Young measure flow as a model for damage

Marc Oliver Rieger and Johannes Zimmer

April 2, 2008

Abstract

Models for hysteresis in continuum mechanics are studied that rely on a time-discretised quasi-static evolution of Young measures akin to a gradient flow. The main feature of this approach is that it allows for local, rather than global minimisation. In particular, the case of a non-coercive elastic energy density of Lennard-Jones type is investigated. The approach is used to describe the formation of damage in a material; existence results are proved, as well as several results highlighting the qualitative behaviour of solutions. Connections are made to recent variational models for fracture.

Keywords Young measures, varifolds, damage, fracture, gradient flows

AMS subject classification numbers 49M20, 74R, 74B20.

1 Introduction

We propose and investigate a simple model of elastic-inelastic material behaviour. Let us sketch the model and focus first on elastic deformations. Their analysis is a well-established field. A classical mathematical description is based on the minimisation of an elastic energy

$$E(u) = \int_\Omega \phi(Du(x))\,\mathrm{d}x, \tag{1}$$

where $\Omega \subset \mathbb{R}^n$ describes the reference configuration of the body under consideration, Du is the deformation gradient, and ϕ is the elastic energy density. To prove existence, it is common to assume coercivity of the functional, that is, a growth condition is imposed on the energy density.

A model of elasticity (like, e.g., classical linear elasticity) can only be assumed to be valid for a finite range of deformations. It seems plausible that, to describe inelastic effects for large deformations (beyond the elastic regime), one has to assume a different growth of the energy density ϕ. Convex-concave energy densities akin to the Lennard-Jones energy $-u_x^{-6} + u_x^{-12}$ have been studied in the engineering literature and related to fracture [37]. They are the *first key ingredient* for the model under consideration. We consider sublinear energy densities as plotted schematically in Figure 1. However, for fracture, Truskinovsky [37] pointed out that the functional (1) with a sublinear energy density such as $\phi(u_x) = \log\left(1 + |u_x(x)|^2\right)$ has a global minimiser with zero energy (for this choice of ϕ; this can be seen by observing that the convex envelope of ϕ is the zero function). The convex-concave energy can be interpreted as a two-well energy with the second well at Infinity. That is, an elastic bar described by this model would break instantaneously if the energy were to obtain its global minimum. A natural strategy is therefore to search for local minimisers. The study of local minimisers (in a suitable topology) is in its infancy; yet, a better understanding may have a number of implications. For example, it has been observed for martensitic phase transitions that dynamics can play an important rôle, preventing the material from attaining the global ground state [20]. This is *the second key ingredient* for the model under consideration: the basis is local minimisation rather than global minimisation.

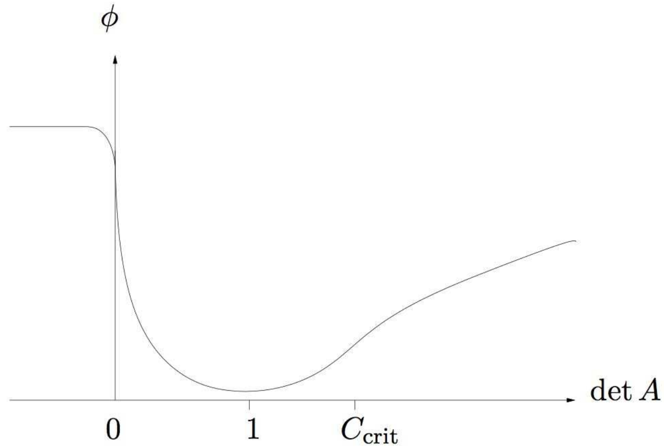

Figure 1: A typical example of an admissible energy density ϕ.

Specifically, we investigate time-discretised versions of a quasi-static gradient flow as a phenomenological model of local energy minimisation for an elastic energy density with sublinear growth at Infinity. It is shown that for initial data in the region representing elastic deformations, stability occurs in the sense that for sufficiently small time steps the response will be elastic (Theorem 3.6 and Theorem 4.1, first part). Instability associated with damage or generally large deformations in the concave region of the energy density is investigated in Theorem 3.8 and Theorem 4.1, second part. It is shown that the proposed time-discretised models exhibit reversible elastic behaviour for small deformations and irreversible effects associated with damage or fracture after increasing the deformation beyond a certain threshold.

A *third ingredient* of the model is the use of Young measure varifolds to describe deformation gradients. Gradient Young measures have been successfully used to deal with nonconvex minimisation problems, in particular for crystallographic microstructures, as pioneered by Ball and James [5] (see, e.g., [30]). Young measure varifolds have been suggested as a generalisation to cover additional concentration effects. We employ Young measure varifolds to describe damage; they provide a unified framework for nonconvex variational problems modelling microstructure, as well as for non-coercive problems describing damage and fracture. Appendix A gives a short synopsis of Young measures and Young measure varifolds.

The proposed models are phenomenological and ignore many effects, such as the precise nature of the influence of defects. One application is the behaviour of an ideal single crystal undergoing potentially arbitrarily large deformations. We believe that a discussion of a simplified model exhibiting some key features of an elastic-inelastic material based on local minimisation can provide valuable insights, while the fundamental ideas can be laid out as clearly as possible. The application of the ideas presented here to fracture will be an area of future research. Here, we focus on the conceptually easier problem of fatigue or material damage. Damage, unlike fracture, is not restricted to lower-dimensional sets.

Obviously, the choice of the topology influences the position of local minimisers and is thus part of the modelling process. To lay out the main ideas without many technicalities, we make a simple choice. We start with a natural metric on the space of probability measures, namely the Wasserstein metric. However, symmetric metrics are found unsuitable in this case. We thus introduce in Section 2 an asymmetric regularisation.

Similar ideas have been used for rate-independent elastoplasticity; Mielke and coworkers have pioneered the mathematical analysis of a derivative-free energetic formulation using Young measures and Wasserstein metrics [26, 27]. The model investigated here is also related to a theory developed by Del Piero, Owen and coworkers [14], who introduced the concept of structured deformations. The fundamental connections of (and differences between) structured deformations and the model introduced here are discussed in Section 5.

We consider two formal discretisations of a quasistatic gradient flow of the energy E with respect to the Wasserstein metric augmented by an asymmetric regularisation. The first discretisation is a standard one, analysed in Section 3. The second discretisation is inspired by De Giorgi's minimising movements [2]; it is investigated in Section 4.

Related models for fracture are discussed in more detail in Section 5. In particular, we discuss the connection to structured deformations and to the Γ-limit of a model proposed by Francfort, Dal Maso and Toader [19, 12]. We close with a discussion in Section 6.

2 Model and mathematical framework

The line of thoughts described in the introduction leads us to a study of local minimisers of non-coercive energy functionals by means of an evolution of Young measures inspired by gradient flows, as suggested by Rieger [33]. Young measure varifolds are only necessary to describe the limiting behaviour as time goes to Infinity. A sketch of such an evolution in the one-dimensional case, which summarises the central idea underlying this article, is depicted in Figure 2. The goal is to study the quasistatic limit of this evolution which can halt at local minimisers, and thus describe hysteretic effects.

We now introduce some notational conventions. In this article, $\Omega \subset \mathbb{R}^n$ is always a bounded (open) domain with smooth boundary. $C_0(\Omega)$ stands for the space of continuous functions $\phi\colon \Omega \to \mathbb{R}$ such that $\{x \in \Omega \mid |\phi(x)| \geq \epsilon\}$ is compact for every $\epsilon > 0$. The essential supremum is denoted sup. For a locally compact Hausdorff space X, we denote the non-negative Radon measures with finite mass by $\mathcal{M}(X)$, and $\mathrm{Prob}(X)$ is the set of probability measures. When no confusion can arise, X is sometimes suppressed from the notation. Further, $\mathcal{M}^p$ denotes, for $1 \leq p \leq \infty$, the set of Radon measures with finite pth moment,

$$\mathcal{M}^p(X) := \{\mu \in \mathcal{M}(X) \mid \int_X |F|^p \,\mathrm{d}\mu(F) < \infty\} \tag{2}$$

The corresponding subset of probability measures with finite pth moment is denoted $\mathrm{Prob}^p(X)$. A sequence $\{\mu_j\}_{j\in\mathbb{N}}$ of probability measures is *tight* if for each $\varepsilon > 0$, there exists a compact set K such that $\mu_j(K) > 1 - \varepsilon$ for every $j \in \mathbb{N}$. The *support* of a Borel measure μ is the complement of the largest open set N with $\mu(N) = 0$. For a given measure μ on X and a given Borel set $B \subset X$, we use the symbol $\mu \llcorner B$ to denote the restriction of the measure μ to B. Weak-$\star$ convergence of a sequence $\{\mu_j\}_{j\in\mathbb{N}}$ of Radon measures to μ holds if $\int_X \phi(x)\,\mathrm{d}\mu_j(x) \to \int_X \phi(x)\,\mathrm{d}\mu(x)$ for every (real-valued) $\phi \in C_0(X)$; this is denoted $\mu_j \overset{\star}{\rightharpoonup} \mu$. We also encounter the concept of narrow convergence [4, Section 5.1], which is defined as follows: a sequence $\{\mu_j\}_{j\in\mathbb{N}}$ of Borel probability measures is said to converge *narrowly* to the Borel probability measure μ as $j \to \infty$ if

$$\lim_{j\to\infty} \int_X \phi(x)\,\mathrm{d}\mu_j(x) = \int_X \phi(x)\,\mathrm{d}\mu(x) \tag{3}$$

for every bounded and continuous (real-valued) function ϕ defined on X.

2.1 Gradient flows and their discretisations

We briefly recollect the mathematical framework for gradient flows. The classical notion of a gradient flow in a Hilbert space,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\nu(t) &= -D\phi(\nu), \\ \nu(0) &= \nu^0, \end{aligned} \tag{4}$$

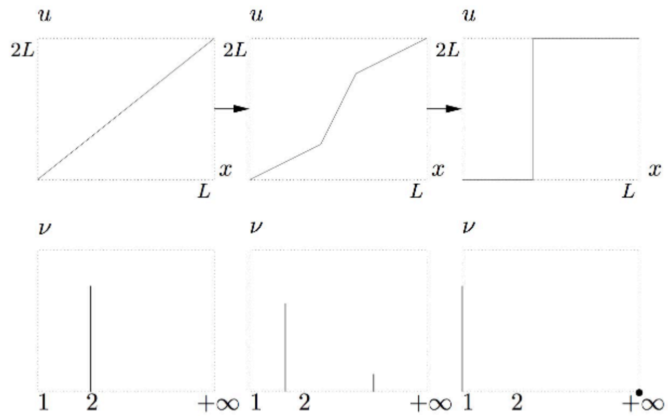

Figure 2: Typical evolution in time (from left to right). Top panel: The initial state has a continuous displacement (top left). The same applies for some intermediate states (top centre), but eventually discontinuities can develop (top right). This corresponds to the occurrence of damage in the material. Bottom panel: the Young measures ν corresponding to the plots in the top panel.

can be generalised to a metric setting [4]. Namely, a gradient flow can be defined by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(\phi \circ \nu) &= -\frac{1}{2}\left|\frac{\mathrm{d}}{\mathrm{d}t}\nu\right|^2 - \frac{1}{2}\left|(D\phi)\circ\nu\right|^2, \\ \nu(0) &= \nu^0, \end{aligned} \tag{5}$$

where $|(D\phi)\circ\nu|$ is an upper gradient [4]. Gradient flows for probability measures can be formulated in this setting; see [4, Part II]. We investigate here a time-discretised version for Young measures (see Appendix A for a brief summary). The set of gradient Young measures with finite first moment (as described in Appendix A) is denoted by $\mathcal{G}$. For simplicity we usually refer to them as *Young measures*.

The motivation for our restriction to time-discretised models is twofold: (i) modelling considerations lead us to the use of an asymmetric metric (see Subsection 2.2 for an explanation). The continuous framework seems so far only to be established for the symmetric case, though an asymmetric version is in preparation [9]. (ii) More importantly, we are interested in *qualitative* results describing the evolution, such as (in-)stability; it seems natural to derive them first in a time-discretised setting. For example, for time-continuous systems of (generalised) Young measures, the control of the correlation of the oscillations has been mastered only recently (see [11], where also the notion of a time derivative is given).

A time-discretised version of the quasi-static version of (4) is

$$\varepsilon\frac{\nu^{j-1}-\nu}{h} = -D\langle\phi,\nu\rangle; \tag{6}$$

the variational model is obtained by regarding (6) as Euler-Lagrange equation. Let us denote the metric by d; then one formally obtains the following model. For a given sequence of deformation gradients $\{A_j\}_{j\in\mathbb{N}}$ with $A_j \in \mathbb{R}^{n\times n}$, let

$$\begin{aligned} X_j := \{\nu \in \mathcal{G} \mid \langle \mathrm{Id}, \nu(x)\rangle = Du(x) \text{ for a.e. } x \in \Omega \\ \text{with } u \in W^{1,\infty}(\Omega,\mathbb{R}^n), u(x) = A_j x \text{ on } \partial\Omega\}. \end{aligned} \tag{7}$$

For a given initial condition $\nu^0 \in X_0$, define the state of the system at discrete time steps $j =$

$1, 2, \ldots$ as a solution of

$$\inf_{\nu \in X_j} \int_\Omega \left[\frac{1}{2} d\left(\nu^{j-1}, \nu\right)^2 + \frac{h}{\varepsilon} \langle \phi, \nu \rangle \right] \mathrm{d}x \tag{8}$$

if such a solution exists. One modification of (8) will be introduced below to arrive at the final model (see (11)); this is to deal with the effects of asymmetry mentioned above. A different approach is presented in Section 4.

2.2 Choice of the metric structure

The choice of the metric structure is part of the modelling process. For the problem under consideration, no particular metric seems to be justified by physical evidence. We use a Wasserstein metric, augmented by an asymmetric metric. The rationale for introducing an asymmetric metric is detailed below. We justify the choice *a posteriori* by proving that essential features of (in)-stability are captured. Other choices of the metric framework are possible and might lead to similar results.

For two (Borel) probability measures ν_1, ν_2 defined on $\mathbb{R}^n$, the *p-Wasserstein metric* is, for $p \in [1, \infty)$, given by

$$d_W^p(\nu_1, \nu_2) := \inf_{T \in \mathrm{Prob}(\nu_1, \nu_2)} \left(\int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^p \, \mathrm{d}T(x, y) \right)^{\frac{1}{p}}, \tag{9}$$

where $\mathrm{Prob}(\nu_1, \nu_2)$ is the set of all probability measures T on $\mathbb{R}^n \times \mathbb{R}^n$ such that $\pi_1 T := \int_{\mathbb{R}^n} \mathrm{d}T(x, \cdot) = \nu_1$ and $\pi_2 T := \int_{\mathbb{R}^n} \mathrm{d}T(\cdot, y) = \nu_2$. The measure T is called a *transport plan* [3]. The Wasserstein metric is sequentially lower semicontinuous (on the set $\mathrm{Prob}^p(\mathbb{R}^n)$ of probability measures with finite pth moment, in the narrow convergence [4, Proposition 7.1.3]; see (3) for the notion of narrow convergence). Here, the Wasserstein metric is restricted to the set $\mathcal{G}$ of gradient Young measures as defined in Appendix A.

For the time-discretisations of a gradient flow model under consideration, the choice $p = 1$ for the Wasserstein metric can be shown to have an appropriate scaling (see [33] for further information). Yet, for a nonconvex energy, the 1-Wasserstein metric alone does not prevent a discontinuous evolution from a situation as in the left panels of Figure 2 directly to the right panels, without intermediate stages [33]. Thus, the choice of the 1-Wasserstein metric alone as a metric would not lead to the existence of the non-global local minimisers we are interested in. We thus augment this metric, and explain now why we choose an asymmetric metric. The metric measures the difference between the Young measures at the previous and the present step. If the metric is symmetric, then a penalty for the nucleation of phases also penalises the disappearance of a phase. To avoid this unphysical behaviour, it seems natural to introduce an asymmetric metric. In this article, we study a special case suitable for our application. Namely, for two bounded sets $A, B \subset \mathbb{R}^{n \times n}$, we define the *upper Hausdorff hemimetric*

$$H_D^+(A; B) := \sup_{b \in B} \inf_{a \in A} d(a, b), \tag{10}$$

where d denotes the Euclidean metric on $\mathbb{R}^{n \times n}$. It might be worthwhile to compare this with the Hausdorff metric

$$H_D(A, B) = \max\left(\sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(a, b) \right).$$

For the model of material damage discussed in this paper, A will be the support of a current phase and B the corresponding support at the next time step; connected components of the support represent phases. Thus, for growing support, B contains A, and hence $H_D^+(A; B)$ and $H_D(A, B)$ agree. For shrinking support, however, B is strictly contained in A, and $H_D^+(A; B) = 0$, whereas the Hausdorff metric is positive. The behaviour of H_D^+ corresponds to the asymmetry in the physical behaviour mentioned above. For two Young measures α and β, the notation

$$d_H^+(\alpha, \beta) := H_D^+(\mathrm{supp}(\alpha); \mathrm{supp}(\beta))$$

is frequently used later on.

In the following, we consider the sum of the 1-Wasserstein metric and a contribution of d_H^+ as follows:

$$\text{Minimise } E\left(\nu^j,\nu\right) := \int_\Omega \left[d_W^1\left(\nu^j,\nu\right)^2 + \frac{h}{\varepsilon}\langle\phi,\nu\rangle\right]\,\mathrm{d}x + \delta^2 \sup_{x\in\Omega} d_H^+\left(\nu^j,\nu\right)^2 \tag{11}$$

among all $\nu \in X_{j+1}$ (again with a given initial condition $\nu^0 \in X_0$).

Let us briefly summarise the model described by (11). The potential energy is given by $\langle\phi,\nu\rangle$, which captures the elastic behaviour. The Young measure ν records the possible formation of microstructures. The dissipation of the process is described by the metric contributions in (11). The asymmetric metric penalises the formation of new phases, while the disappearance of phases is not penalised. Phases correspond to connected components of the support of ν.

The model (11) is unusual in the sense that the asymmetric contribution $\delta^2 \sup_{x\in\Omega} d_H^+\left(\nu^j,\nu\right)^2$ is non-local. Nonlocal models, however, have been investigated at least since Eringen's work in the 1960s; we refer to the article by Chen, Lee and Eskandarian [8] for a recent survey. Peridynamics is another example of a non-local continuum theory [36, 24]. We employ the L^∞-norm for the Hausdorff term, rather than an L^p-norm with $1 \le p < \infty$. To demonstrate the difficulties in an L^p-setting with $p < \infty$, we wish to show that transport over relatively large distances is then not excluded, which could lead to an instability of a local minimum and might allow for the instantaneous formation of fracture. This is demonstrated with the following one-dimensional example, where we restrict for simplicity the class of admissable functions to Dirac measures. Take $\Omega := (0,1)$, $A_j = 2$ for $j = 0,1,\ldots$ and $u_0(x) = 2x$, $\nu_0(x) = \delta_2$. Moreover, choose ϕ such that $\phi(2) = 1$, $\phi'(2) = 6$ and ϕ strictly convex in the region under consideration, e.g., on $\left[\frac{1}{2},3\right]$. If we assume $p < \infty$, then, for $\alpha,\varepsilon > 0$,

$$\nu_1(x) := \begin{cases} \delta_{2+\alpha+\varepsilon^2} & \text{for } x \ge \varepsilon, \\ \delta_{\frac{\alpha}{\varepsilon}+\varepsilon^2} & \text{for } x < \varepsilon \end{cases}$$

has from ν_0 the Hausdorff (semi)-distance

$$\int_0^1 \left|d_H^+\left(\nu_0(x),\nu_1(x)\right)\right|^p\,\mathrm{d}x = (1-\varepsilon)\left(\alpha+\varepsilon^2\right)^p + \varepsilon\left(2-\frac{\alpha}{\varepsilon}-\varepsilon^2\right)^p.$$

If we set $\alpha = \varepsilon$, the expectation value of ν_1 is again 2. Moreover we can estimate further for $\varepsilon < 1$

$$(1-\varepsilon)\left(\alpha+\varepsilon^2\right)^p + \varepsilon\left(2-\frac{\alpha}{\varepsilon}-\varepsilon^2\right)^p = (1-\varepsilon)\varepsilon^p + \varepsilon + \mathcal{O}\left(\varepsilon^2\right) \le 2\varepsilon + \mathcal{O}\left(\varepsilon^2\right). \tag{12}$$

The crucial observation is that this term is now an order of magnitude smaller than it would have been in the L^∞-setting we have chosen. The Wasserstein part becomes

$$\int_\Omega d_W^1\left(\nu_0,\nu_1\right)^2\,\mathrm{d}x \le 2\varepsilon + \mathcal{O}\left(\varepsilon^2\right). \tag{13}$$

On the other hand, the elastic energy of ν_1 is reduced by the amount

$$\begin{aligned}\int_0^1 \left[\langle\phi,\nu_0\rangle - \langle\phi,\nu_1\rangle\right]\,\mathrm{d}x &= -(1-\varepsilon)\phi'(2)\left(\alpha+\varepsilon^2\right) + \mathcal{O}\left(\left(\alpha+\varepsilon^2\right)^2\right) + \varepsilon\left(\phi\left(\frac{\alpha}{\varepsilon}+\varepsilon^2\right)-\phi(2)\right)\\ &= -\varepsilon\left[(1-\varepsilon)\phi'(2) - \phi(2) + \mathcal{O}(\varepsilon)\right]\\ &= -\varepsilon\left[6(1-\varepsilon) - 1 + \mathcal{O}(\varepsilon)\right]\\ &= -5\varepsilon + \mathcal{O}\left(\varepsilon^2\right). \end{aligned} \tag{14}$$

For $\varepsilon > 0$ sufficiently small, the sum of the three terms (12)–(14) is negative. This indicates that the L^p-version of the metric d_H^+ is not suitable to prevent a transport of mass on a small set in Ω over large distances. It is likely that a local minimum in the elastic region may be unstable; an immediate onset of damage on a small set of measure $\varepsilon > 0$ may follow. In the case $p = \infty$, this problem cannot occur, as we will see.

The mathematical properties of (11) are studied in Section 3; an alternative approximation in the same metric setting is studied in Section 4. We now collect some auxiliary results on the hemimetric H_D^+.

2.3 Auxiliary results on the upper Hausdorff hemimetric

We state some facts on the upper Hausdorff hemimetric H_D^+ defined in (10).

Lemma 2.1 *Let $A, B, C \subset \mathbb{R}^{n\times n}$ be bounded sets. The upper Hausdorff hemimetric H_D^+ can be equivalently characterised as $H_D^+(A; B) = \inf\{r > 0 \mid B \subset N_r(A)\}$, where $N_r(A) := \cup_{a\in A} B(a, r)$. Furthermore, H_D^+ satisfies $H_D^+(A; B) \geq 0$, $H_D^+(A; B) = 0$ if and only if $\bar{A} \supset B$, and $H_D^+(A; C) \leq H_D^+(A; B) + H_D^+(B; C)$.*
If $B \setminus A \neq \emptyset$ then

$$H_D^+(A; B) = \sup_{b\in B\setminus A} \inf_{a\in A} d(a, b). \tag{15}$$

Proof: All claims are easy to verify, and we only give the proof of the triangle inequality. There exist points $a \in \overline{A}$ and $c \in \overline{C}$ such that $d(a, c) = H_D^+(A; C)$. There exists $b \in \overline{B}$ such that $d(b, c) \leq H_D^+(B; C)$. For this b, one can choose $\tilde{a} \in \overline{A}$ such that $d(\tilde{a}, b) \leq H_D^+(A; B)$. Then $H_D^+(A; C) = d(a, c) \leq d(\tilde{a}, c) \leq d(\tilde{a}, b) + d(b, c) \leq H_D^+(A; B) + H_D^+(B; C)$. □

We need to discuss the convergence of probability measures, and the right notion is that of weak-$\star$ convergence. We refer the reader to the beginning of Section 2 for the definition.

Lemma 2.2 *Let $\{\nu_j\}_{j\in\mathbb{N}}$ be a sequence of probability measures with $\nu_j \overset{\star}{\rightharpoonup} \nu$ as $j \to \infty$. Then for every point $F \in supp(\nu)$ there is a sequence of points $F_j \in supp(\nu_j)$ such that $F_j \to F$.*

Proof: Suppose the contrary. Then there is an open neighbourhood U of F such that $U \cap \operatorname{supp}(\nu_j) = \emptyset$ for every $j \in \mathbb{N}$ sufficiently large. As a consequence of the weak-$\star$ convergence [17, Chapter 1, Theorem 3], we have $0 \leq \nu(U) \leq \liminf_{j\to\infty} \nu_j(U) = 0$. Thus, $\nu(U) = 0$, and hence $F \notin \operatorname{supp}(\nu)$ contrary to the initial assumption. □

Besides the physical motivation outlined above, there is also a mathematical reason for introducing d_H^+. Namely, d_H^+ is the weak-$\star$ sequential lower semicontinuous envelope of the Hausdorff metric with respect to the second argument.

Lemma 2.3 *Let α and β be probability measures; let $d_H(\alpha, \beta) := H_D(supp(\alpha), supp(\beta))$ and let $\widetilde{d_H}$ be its weak-$\star$ sequential lower semicontinuous envelope with respect to its second variable. Then $\widetilde{d_H}(\alpha, \beta) = d_H^+(\alpha, \beta)$.*

Proof: We consider a sequence of probability measures $\{\beta_n\}_{n\in\mathbb{N}}$ with $\beta_n \overset{\star}{\rightharpoonup} \beta$ as $n \to \infty$. Then, with $A := \operatorname{supp}(\alpha)$ and $B_n := \operatorname{supp}(\beta_n)$, we estimate $\widetilde{d_H}$ from below as follows (the second inequality uses Lemma 2.2).

$$\begin{aligned}
\widetilde{d_H}(\alpha, \beta) &= \inf_{\beta_n \overset{\star}{\rightharpoonup} \beta} \liminf_{n\to\infty} H_D(\operatorname{supp}(\alpha), \operatorname{supp}(\beta_n)) \\
&= \inf_{\beta_n \overset{\star}{\rightharpoonup} \beta} \liminf_{n\to\infty} \max\left[\sup_{a\in A} \inf_{b\in B_n} d(a, b), \sup_{b\in B_n} \inf_{a\in A} d(a, b)\right] \\
&\geq \inf_{\beta_n \overset{\star}{\rightharpoonup} \beta} \liminf_{n\to\infty} \sup_{b\in B_n} \inf_{a\in A} d(a, b) \\
&= \inf_{\beta_n \overset{\star}{\rightharpoonup} \beta} \liminf_{n\to\infty} H_D^+(\operatorname{supp}(\alpha); \operatorname{supp}(\beta_n)) \\
&\geq H_D^+(\operatorname{supp}(\alpha); \operatorname{supp}(\beta)) \\
&= d_H^+(\alpha, \beta).
\end{aligned}$$

For $\beta_n := \left(1-\frac{1}{n}\right)\beta + \frac{1}{n}\alpha$, it is easy to see that $\sup_{a\in A}\inf_{b\in B_n} d(a,b) = 0$ and $\beta_n \overset{\star}{\rightharpoonup} \beta$ as $n\to\infty$. This shows that equality can be obtained in the preceding estimate. $\square$

2.4 Assumptions on the energy density

For the mathematical analysis of the two time-discretised models proposed here, we need to make assumptions on the sublinear energy density $\phi\colon \mathbb{R}^{n\times n}\to\mathbb{R}$. We consider energies akin to the Lennard-Jones energy $\phi(r) := \frac{1}{r^{12}} - \frac{1}{r^6}$ in the sense that the energy density allows for damage by penalising large deformation gradients in a suitable way (not ruling out ever larger deformation gradients). To be specific, we collect the following assumptions on ϕ (see Appendix A for the definition of (uniform) quasiconvexity):

Assumption 2.4 (Energy density function) *We assume that the energy density $\phi\colon \mathbb{R}^{n\times n}\to \mathbb{R}$ is a function of the deformation gradient and satisfies the following conditions:*

(i) $\phi\in C^2(\mathbb{R}^{n\times n},\mathbb{R})$, *$\phi(RA)=\phi(A)$ for every $R\in \mathrm{SO}(n)$.*

(ii) There exists a positive constant C_{crit} and a uniformly quasiconvex function $\tilde{\phi}\colon \mathbb{R}^{n\times n}\to\mathbb{R}$ such that for $\mathcal{R} := \{A\in\mathbb{R}^{n\times n} \mid 0\le \det(A) < C_{\mathrm{crit}}\}$, the equality $\tilde{\phi}\big|_{\mathcal{R}} = \phi\big|_{\mathcal{R}}$ holds.

(iii) For every $A\in\mathbb{R}^{n\times n}$ with $\det(A)=1$, the function $\lambda\mapsto\phi(\lambda A)$ takes its minimum at $\lambda=1$ where $\phi(A)=0$. Moreover, this function is strictly increasing on $(1,C_{\mathrm{crit}})$.

(iv) For every $A\in\mathbb{R}^{n\times n}$ with $\det(A)>C_{\mathrm{crit}}$ there exists a rank-one matrix C such that $\lambda\mapsto \phi(A+\lambda C)$ is strictly concave at $\lambda=0$.

(v) ϕ is sublinear at Infinity, that is, there exists a constant $C_1>0$ such that $\phi(A)\le C_1\det(A)$ for every $A\in\mathbb{R}^{n\times n}$ with $\det(A)>1$.

Remark 2.5 *1. One can easily check that for $n=1$, all conditions of Assumption 2.4 are satisfied, e.g., by a suitably scaled and shifted energy of Lennard-Jones type (modified to attain a large, but finite constant for negative values of A, compare Figure 1). The energy we have in mind is of this type (with $\det(A)$ as argument if $n>1$). To give an example, let us consider*

$$\tilde{\phi}(A) := \begin{cases} (\det(A)-1)^2 & \text{for } \det(A)\in(0,C_{\mathrm{crit}}),\\ 1 & \text{for } \det(A)\le 0,\\ \psi(|\det(A)|) & \text{for } \det(A)\ge C_{\mathrm{crit}}, \end{cases}$$

where $\psi\colon [C_{\mathrm{crit}},\infty)\to\mathbb{R}$ is a sublinear, increasing function satisfying the compatibility assumptions $\psi(C_{\mathrm{crit}}) = (C_{\mathrm{crit}}-1)^2$, $\psi'(C_{\mathrm{crit}}) = 2C_{\mathrm{crit}}$ and $\psi''(C_{\mathrm{crit}})=2$. To meet the regularity assumption (i), ϕ needs to be smoothened in a neighbourhood of zero; a suitable mollification ϕ of $\tilde{\phi}$ then meets all assumptions. The energy density ϕ constructed in this way is purely volumetric. However, even if the energy ϕ is reminiscent of that of a Lennard-Jones fluid, it is important to recognise that the models studied in Section 3 and 4 have a penalisation built in for the formation of arbitrarily large shears, namely via the term with the Wasserstein metric for Young measures. This term measures the distance to the state at the previous time-step; Equation (38) in Appendix A shows that for classical functions, this contribution amounts to a regular L^1-contribution to the energy. Thus the deviatoric evolution is controlled via the Wasserstein term.

2. Some of our results in Section 3 are only proved for the one-dimensional case $n=1$.

3. The conditions (ii)–(v) in Assumption 2.4 are not required for the existence result in Theorem 3.2.

4. *The finiteness of ϕ is a technical assumption for the existence results (Theorem 3.2 and Theorem 4.1). Since the behaviour of the material under compression is not the focus of this paper, we refrain from weakening the assumption on ϕ for matrices with negative determinant.*

A typical example of a one-dimensional admissible energy density is shown in Figure 1.

3 Existence and stability results in the standard discretisation

In this section, we prove existence of a solution to the variational problem (11) in arbitrary space dimensions and study its evolution in a one-dimensional setting. We work with Young measures and their strong topology; for a synopsis of Young measures and their topologies, we refer to Appendix A.

3.1 Existence results

Lemma 3.1 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with smooth boundary. Let $\nu \in \mathcal{G}$, that is, ν is a $W^{1,\infty}$-gradient Young measure defined on Ω. Let supp$(\nu(x))$ be compact for every $x \in \Omega$. Then $\sup_x d_H^+(\nu(x), \cdot)$ is sequentially lower semicontinuous on $\mathcal{G}$ in the weak⋆-topology.*

Proof: This follows immediately from Lemma 2.3. Namely, for fixed $x \in \Omega$ and $\mu_j(x) \overset{\star}{\rightharpoonup} \mu(x)$ as $j \to \infty$,

$$d_H^+(\nu(x), \mu(x)) \leq \liminf_{j\to\infty} d_H^+(\nu(x), \mu_j(x)) \leq \liminf_{j\to\infty} \sup_{x\in\Omega} d_H^+(\nu(x), \mu_j(x)),$$

and the claim follows by taking the supremum over $x \in \Omega$ on both sides. □

Theorem 3.2 *For every $h, \varepsilon, \delta > 0$ and for $\nu^j \in \mathcal{G}$ with supp$\left(\nu^j(x)\right)$ bounded uniformly in $x \in \Omega$, the variational problem (11) admits a solution $\nu \in \mathcal{G}$.*

Proof: This follows by the direct method from the calculus of variations. Namely, $\mathcal{G}$ is a closed subset of the Banach space of Radon measures on $\Omega \times \mathbb{R}^{n\times n}$; for fixed $\nu^j \in X_j$, the functional

$$\nu \mapsto E(\nu^j, \nu) := \int_\Omega \left[d_W^1\left(\nu^j, \nu\right)^2 + \frac{h}{\varepsilon}\langle \phi, \nu\rangle \right] \mathrm{d}x + \delta^2 \sup_{x\in\Omega} d_H^+\left(\nu^j, \nu\right)^2$$

is bounded from below. Hence there exists a minimising sequence $\{\nu_k\}_{k\in\mathbb{N}}$ of gradient Young measures ν_k for E. Since $\|\nu_k(x)\| = 1$ for a.e. $x \in \Omega$, the whole measures are uniformly bounded, $\|\nu_k\| = |\Omega| < \infty$. Furthermore, $d_W^1\left(\nu^j, \nu_k\right)^2$ is bounded uniformly in k. Since the space of Young measures with the 1-Wasserstein metric (37) is complete, there exists a subsequence (not relabeled) and a Young measure ν with $\nu_k \to \nu$ in the Wasserstein topology. Since convergence in the Wasserstein topology implies weak⋆-convergence, we find $\nu_k \overset{\star}{\rightharpoonup} \nu$ (in the sense of measures).

To show that the limit measure ν is a gradient Young measure, consider for every ν_k a sequence $\{f_{k,l}\}_{l\in\mathbb{N}}$ of gradients converging to ν_k (in the sense of Young measures). Then a diagonal argument shows that ν is also a gradient Young measure, hence $\nu \in \mathcal{G}$.

Finally, E is sequential lower semicontinuous in the Wasserstein topology. For the regularising Hausdorff term $\sup_x d_H^+(\nu(x), \cdot)$, this follows from Lemma 3.1, since the Wasserstein metric metrises weak⋆-convergence [38, Theorem 7.12]. The lower semicontinuity of the other terms is immediate. In summary, the limit measure ν minimises the problem (11). (We remark that for spatially homogeneous Young measures, one can argue differently. Namely, recall the notion of narrow convergence (3); the Hausdorff term yields narrow convergence of the minimising sequence. This combined with Fatou's lemma gives the result. See also [22, Proof of Proposition 4.1] for the sequential lower semicontinuity of the Wasserstein metric.) □

3.2 Stability and instability

Now that we have established existence, our next goal is to study qualitative properties of the solutions in the one-dimensional case $\Omega \subset \mathbb{R}$. We return to the general case $\Omega \subset \mathbb{R}^n$ with $n \geq 1$ in Section 4. We show that the proposed model exhibits a behaviour akin to damage in a material. The central feature of the model is that below a certain deformation threshold, processes are reversible, i.e., the material behaves elastically. Above the threshold, irreversibility (damage) occurs. The model combines both behaviours in a unified perspective via a single variational principle. To simplify the argumentation, we consider the spatially homogeneous situation by assuming that all measures are constant in $x \in \Omega$. We also assume that the expectation value A_j is independent of $j \in \mathbb{N}$. We denote its value by A and write $X := X_j$. We are only interested in bounded, connected domains Ω and may assume without loss of generality that $\Omega := (0,1)$. We also write "the convex region of ϕ" for the set of all $A \in \mathbb{R}$ with $0 < A < C_{\mathrm{crit}}$ (compare Assumption 2.4 (ii)). Some of our results have been proved previously [33] for the simpler case of measures concentrated in exactly two points.

We introduce some terminology and an auxiliary result:

Definition 3.3 (Submeasure) *Let β be a nonnegative Radon measure on $\mathbb{R}^n$. A nonnegative measure α defined on $\mathbb{R}^n$ is a* submeasure *of β if $\alpha(B) \leq \beta(B)$ for all measurable sets $B \subset \mathbb{R}^n$. We then write $\alpha \leq \beta$.*

Lemma 3.4 (Generalised Jensen Inequality) *Let $\phi\colon \mathbb{R} \to \mathbb{R}$ be a strictly convex function. Let $a \leq b$ and σ, σ' be two non-negative measures on $\mathbb{R}$ with $\sigma \neq \sigma'$ and $\|\sigma\| = \|\sigma'\|$, such that supp $(\sigma) \subset [a,b]$ and supp $(\sigma') \cap (a,b) = \emptyset$. Moreover, assume that the expectation value of σ and the expectation value of σ' coincide. Then $\langle \phi, \sigma\rangle < \langle \phi, \sigma'\rangle$.*

The standard Jensen inequality can be deduced as a special case of this result by taking $a = b$. The proof of the generalization is straightforward and is given here for the reader's convenience.

Proof: We want to show

$$\int_{\mathbb{R}} \phi(x)\,\mathrm{d}\sigma(x) < \int_{\mathbb{R}} \phi(x)\,\mathrm{d}\sigma'(x). \tag{16}$$

Since σ and σ' have the same expectation value, we can add any affine function $x \mapsto \lambda x$ to ϕ without changing this inequality. In particular, we can obtain that $\phi(a) + \lambda a = \phi(b) + \lambda b$, if $\phi(a) \neq \phi(b)$. We define $\tilde{\phi}(x) := \phi(x) + \lambda x$. Since then $c := \tilde{\phi}(a) = \tilde{\phi}(b)$ by construction, and $\tilde{\phi}$ is less or equal than c in (a,b) by convexity of $\tilde{\phi}$, where the support of σ is located, but larger than c outside $[a,b]$, where the support of σ' is located (again by convexity of $\tilde{\phi}$), it is now easy to prove (16) with $\tilde{\phi}$ instead of ϕ:

$$\int_{\mathbb{R}} \tilde{\phi}(x)\,\mathrm{d}\sigma(x) = \int_a^b \tilde{\phi}(x)\,\mathrm{d}\sigma(x) \leq c\,\|\sigma\| = c\,\|\sigma'\| \leq \int_{\mathbb{R}} \tilde{\phi}(x)\,\mathrm{d}\sigma'(x).$$

Equality can only hold if $\sigma = \sigma'$, hence we have proved the lemma. □

We are now in a position to state some stability results. The first stability result shows that, if the deformation gradients are initially contained in the convex region of ϕ, then the difference between the smallest and largest deformation gradient shrinks over time unless the initial gradient is already concentrated, $\nu^j = \delta_A$.

Lemma 3.5 *If $A_j := A < C_{\mathrm{crit}}$ for every $j \in \mathbb{N}$, and $\nu^j \in \mathcal{G}$ is a homogeneous Young measure supported in $(0, C_{\mathrm{crit}})$, and a solution ν^{j+1} of (11) is also supported in $(0, C_{\mathrm{crit}})$, then*

$$\left(\mathit{supp}\left(\nu^{j+1}\right)\right)^{\mathrm{conv}} \subsetneq \left(\mathit{supp}\left(\nu^{j}\right)\right)^{\mathrm{conv}},$$

unless $\nu^j = \delta_A$, in which case $\nu^{j+1} = \nu^j$.

Proof: Let us define $c^j := \max \operatorname{supp}\left(\nu^j\right)$ and suppose, for contradiction, $c^{j+1} > c^j$. We recall that $\mu \llcorner B$ denotes the restriction of the measure μ to B. Let $T \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$ be a transport plan [3] as described in Subsection 2.2. Specifically, let T be a transport plan that minimises the L^1-Wasserstein metric and has the marginals $\pi_1 T = \nu^j$ and $\pi_2 T = \nu^{j+1}$. Let us consider arbitrary ν^j-measurable sets $A, B \subset \mathbb{R}$ with $\nu^j(A) > 0$ and $\nu^j(B) > 0$ whose support does not overlap in the sense that $\inf\{x \in B\} > \sup\{x \in A\}$ holds. We consider their images under the transport plan T,

$$A' := \operatorname{supp}\left(\pi_2\left(T \llcorner A \times \mathbb{R}\right)\right) \text{ and } B' := \operatorname{supp}\left(\pi_2\left(T \llcorner B \times \mathbb{R}\right)\right).$$

If the images have the same ordering, namely $\inf\{x \in B'\} \geq \sup\{x \in A'\}$, then we say that T is *monotone increasing.* Here, we may assume that T is monotone increasing; the existence of such a monotone transport plan has been established elsewhere [34].

For an arbitrary $\eta \in \left(0, c^{j+1} - c^j\right)$, we define $\mu^{j+1} := \nu^{j+1} \llcorner [c^{j+1} - \eta, c^{j+1}]$. Let us also introduce $\mu^j := \pi_1 T \llcorner \mathbb{R} \times [c^{j+1} - \eta, c^{j+1}]$. We say that T transports μ^j to μ^{j+1}. Since T is monotone, we obtain [34]

$$\inf \operatorname{supp}\left(\mu^j\right) \geq \sup \operatorname{supp}\left(\nu^j - \mu^j\right).$$

We observe that $\inf \operatorname{supp}\left(\mu^{j+1}\right) > \sup \operatorname{supp}\left(\mu^j\right)$ by definition of η and consequently $\langle \mathrm{Id}, \mu^j \rangle < \langle \mathrm{Id}, \mu^{j+1} \rangle$. Furthermore, again by definition, $\langle \mathrm{Id}, \nu^j \rangle = A = \langle \mathrm{Id}, \nu^{j+1} \rangle$. Thus, there exists a submeasure $\tau^{j+1} \leq \nu^{j+1}$ and a submeasure $\tau^j \leq \nu^j$ such that T transports τ^j to τ^{j+1} and $\langle \mathrm{Id}, \tau^{j+1} + \mu^{j+1} \rangle = \langle \mathrm{Id}, \tau^j + \mu^j \rangle$. Hence we have $\langle \mathrm{Id}, \tau^j \rangle > \langle \mathrm{Id}, \tau^{j+1} \rangle$.

Let Σ be the part of T that transports τ^j to τ^{j+1}, that is, $\Sigma \leq T$ with marginals $\pi_1 \Sigma = \tau^j$ and $\pi_2 \Sigma = \tau^{j+1}$. Then, due to the monotonicity of T, the set $S := \operatorname{supp}(\Sigma) \subset \mathbb{R}^2$ is the graph of a monotone correspondence from $\operatorname{supp}\left(\tau^j\right)$ to $\operatorname{supp}\left(\tau^{j+1}\right)$. Since $\langle \mathrm{Id}, \tau^{j+1} \rangle < \langle \mathrm{Id}, \tau^j \rangle$, there must be some set $A_0 \subset \operatorname{supp}\left(\tau^j\right)$ such that $\pi_2\left(T \llcorner \{(A_0, \mathbb{R})\}\right)$ has positive measure on $(-\infty, x_0)$, where $x_0 := \inf A_0$. We define $\hat{\tau}^j$ as a submeasure of τ^j such that $\hat{\tau}^j$ is transported by T on a subset $\hat{\tau}^{j+1}$ of $(-\infty, x_0)$. By this method we still have $\langle \mathrm{Id}, \hat{\tau}^{j+1} \rangle < \langle \mathrm{Id}, \hat{\tau}^j \rangle$. Then there is a submeasure $\hat{\mu}^j$ of μ^j such that, with its image $\hat{\mu}^{j+1}$ under the transport plan,

$$\langle \mathrm{Id}, \hat{\tau}^{j+1} + \hat{\mu}^{j+1} \rangle = \langle \mathrm{Id}, \hat{\tau}^j + \hat{\mu}^j \rangle.$$

We have now neatly separated the supports of $\hat{\mu}^j + \hat{\tau}^j$ and $\hat{\mu}^{j+1} + \hat{\tau}^{j+1}$, since

$$\begin{aligned} \sup \operatorname{supp}\left(\hat{\tau}^{j+1}\right) < x_0 < \inf \operatorname{supp}\left(\hat{\tau}^j\right) \leq \sup \operatorname{supp}\left(\hat{\tau}^j\right) \\ < \inf \operatorname{supp}\left(\hat{\mu}^j\right) \leq \sup \operatorname{supp}\left(\hat{\mu}^j\right) < c^j < \inf \operatorname{supp}\left(\hat{\mu}^{j+1}\right). \end{aligned}$$

Therefore we can apply Lemma 3.4 with $\sigma := \hat{\mu}^j + \hat{\tau}^j$, $\sigma' := \hat{\mu}^{j+1} + \hat{\tau}^{j+1}$, $a := x_0$ and $b := c^j$ to prove that

$$\langle \phi, \hat{\mu}^j + \hat{\tau}^j \rangle < \langle \phi, \hat{\mu}^{j+1} + \hat{\tau}^{j+1} \rangle.$$

We can thus define $\tilde{\nu}^{j+1} := \nu^{j+1} - \tau^{j+1} - \mu^{j+1} + \hat{\tau}^j + \hat{\mu}^j$ and deduce $E\left(\nu^j, \tilde{\nu}^{j+1}\right) < E\left(\nu^j, \nu^{j+1}\right)$. This is a contradiction to the assumption that ν^{j+1} is a minimiser of the time step problem.

In an analogous way one can rule out that $\inf \operatorname{supp}\left(\nu^{j+1}\right) < \inf \operatorname{supp}\left(\nu^j\right)$. This shows that $\left(\operatorname{supp}\left(\nu^{j+1}\right)\right)^{\mathrm{conv}} \subset \left(\operatorname{supp}\left(\nu^j\right)\right)^{\mathrm{conv}}$.

To prove that the convex hull of the support is in fact shrinking, we use a similar idea by defining $b^{j+1} := \inf \operatorname{supp}\left(\nu^{j+1}\right)$, $\mu^{j+1} := \nu^{j+1} \llcorner [c^{j+1} - \eta, c^{j+1}]$ and $\tau^{j+1} := \nu \llcorner [b^{j+1}, b^{j+1} + \eta]$. One can shift μ^{j+1} slightly to the left by setting $\tilde{\mu}^{j+1} := \mu^{j+1}(\cdot + \eta)$, and adjust the mean value accordingly by shifting τ^{j+1} to the right, $\tilde{\tau}^{j+1} := \tau^{j+1}(\cdot - r(\eta))$, where $r(\eta) > 0$ is chosen such that $\langle \mathrm{Id}, \tau^{j+1} + \mu^{j+1} \rangle = \langle \mathrm{Id}, \tilde{\tau}^{j+1} + \tilde{\mu}^{j+1} \rangle$. Obviously, $r(\eta) = O(\eta)$ for $\eta \to 0$. A straightforward estimate shows that for $\tilde{\nu}^{j+1} := \nu^{j+1} - \mu^{j+1} - \tau^{j+1} + \tilde{\mu}^{j+1} + \tilde{\tau}^{j+1}$ and $\eta > 0$, sufficiently small,

$$E(\nu^j, \nu) - E(\nu^j, \tilde{\nu}) \leq -\frac{h}{\varepsilon} \eta \, |\mu| + |\mu|^2 \, \eta^2 + |\tau|^2 \, c(\eta)^2 + \delta^2 \eta^2 + \delta^2 r(\eta)^2.$$

For sufficiently small $\eta > 0$, this becomes negative, which contradicts the assumption that ν^{j+1} is a minimiser of the time step problem. $\square$

The previous result does not exclude concentration (i.e., leaking of mass of ν^j to Infinity) for initial conditions within the convex region of ϕ, since it is *assumed* that the deformation gradient does not leave the convex region of ϕ. In other words, we can so far only prove stability if supp $\left(\nu^j\right)$ is contained in the convex part of ϕ. The next theorem, however, ensures that no damage occurs when the deformation is below a certain threshold C_{crit} defined by the region of convexity of the energy density ϕ if the time-discretisation h is small enough. In other words, one observes elastic behaviour in this region.

Theorem 3.6 (Stability I) *Let $\Omega = (0,1)$, and let $\eta \in (0, C_{\text{crit}})$, and suppose $A_j := A < C_{\text{crit}}$ for every $j \in \mathbb{N}$. Let us assume furthermore that the initial Young measure ν^0 satisfies supp $\left(\nu^0\right) \subset (0, C_{\text{crit}} - \eta]$ for some $\eta > 0$. Then, for h small enough and $\delta^2 = \frac{h}{\varepsilon}$, the solution remains in $(0, C_{\text{crit}} - \eta]$ for all times. Namely, any sequence of solutions ν^j of (11) satisfies supp $\left(\nu^j\right) \subset (0, C_{\text{crit}} - \eta]$ for all $j \in \mathbb{N}$.*

Proof: For fixed $j \in \mathbb{N}$, the minimisation problem (11) on $\Omega = (0,1)$ in the spatially homogeneous case reads

$$E(\nu^j, \nu) = \frac{h}{\varepsilon}\langle \phi, \nu\rangle + d_W^1\left(\nu^j, \nu\right)^2 + \delta^2 d_H^+\left(\nu^j, \nu\right)^2. \tag{17}$$

Denote, for a fixed time step $h > 0$, a minimiser of (17) by ν_h^{j+1}. Then $E\left(\nu^j, \nu_h^{j+1}\right) - E\left(\nu^j, \nu^j\right) \leq 0$ holds, and thus, since $E\left(\nu^j, \nu^j\right) = \frac{h}{\varepsilon}\langle \phi, \nu^j\rangle$,

$$\frac{h}{\varepsilon}\langle \phi, \nu_h^{j+1}\rangle - \frac{h}{\varepsilon}\langle \phi, \nu^j\rangle + d_W^1\left(\nu^j, \nu_h^{j+1}\right)^2 + \delta^2 d_H^+\left(\nu^j, \nu_h^{j+1}\right)^2 \leq 0. \tag{18}$$

We observe that, with $F^j := \inf \operatorname{supp}\left(\nu^j\right)$ and Assumption 2.4 (ii), $\langle \phi, \nu^j\rangle$ is bounded from above by $\max\{\phi(F^j), \phi(C_{\text{crit}})\}$. Since $\langle \phi, \nu_h^{j+1}\rangle$ is bounded from below by Assumption 2.4 (iii), there is a constant $C_1 > 0$ such that

$$\langle \phi, \nu^j\rangle - \langle \phi, \nu_h^{j+1}\rangle \leq C_1. \tag{19}$$

Using (18), we obtain the estimate

$$0 \leq d_W^1\left(\nu^j, \nu_h^{j+1}\right)^2 \leq \frac{h}{\varepsilon}\left(\langle \phi, \nu^j\rangle - \langle \phi, \nu_h^{j+1}\rangle\right) \leq C_1 \frac{h}{\varepsilon}. \tag{20}$$

We would like to use Inequality (20) to improve the Estimate (19). To this end, we employ the fact that the 1-Wasserstein metric is the dual norm of the Lipschitz norm (see, for example, [32, Section 5.3], [25, Section 2] or [26, Subsection 5.2]),

$$d_W^1(\mu, \nu) = \sup_{\operatorname{Lip}(\phi) \leq 1} \langle \phi, \mu - \nu\rangle.$$

This implies immediately for $\mu := \nu^j$ and $\nu := \nu_h^{j+1}$

$$\langle \phi, \nu^j\rangle - \langle \phi, \nu_h^{j+1}\rangle \leq L d_W^1\left(\nu^j, \nu_h^{j+1}\right), \tag{21}$$

where L denotes the Lipschitz constant of ϕ on (F^0, ∞), which is finite due to the regularity and the growth condition of ϕ (Assumption 2.4 (v)).

The combination of (20) and (21) shows that there is a constant $C_2 = C_2\left(\phi, \frac{h}{\varepsilon}\right)$ depending only on the energy density ϕ and $\frac{h}{\varepsilon}$ such that

$$\langle \phi, \nu^j\rangle - \langle \phi, \nu_h^{j+1}\rangle \leq C_2\left(\phi, \tfrac{h}{\varepsilon}\right) \to 0 \text{ as } \frac{h}{\varepsilon} \to 0. \tag{22}$$

This information, combined with (18), yields for $\delta^2 = \frac{h}{\varepsilon}$ as assumed here

$$d_H^+\left(\nu^j, \nu_h^{j+1}\right)^2 \leq \langle \phi, \nu^j\rangle - \langle \phi, \nu_h^{j+1}\rangle \leq C_2\left(\phi, \tfrac{h}{\varepsilon}\right).$$

By (22), we can choose $\frac{h}{\varepsilon}$ small enough that $C_2\left(\phi, \frac{h}{\varepsilon}\right) < \sqrt{\eta}$. Then $\operatorname{supp}\left(\nu^{j+1}\right) \subset (0, C_{\text{crit}})$. We then apply Lemma 3.5 to deduce $\left(\operatorname{supp}\left(\nu^{j+1}\right)\right)^{\text{conv}} \subset \left(\operatorname{supp}\left(\nu^{j}\right)\right)^{\text{conv}}$. Hence $\operatorname{supp}\left(\nu^{j+1}\right) \subset (0, C_{\text{crit}} - \eta]$. Since the choice of $\frac{h}{\varepsilon}$ did not depend on j, one can apply this argument at every time step j, which proves Theorem 3.6. $\square$

We have proved stability for data supported in the convex region of ϕ, that is, for measures supported on $(0, C_{\text{crit}})$, and constant $A < C_{\text{crit}}$. That is, a body under tension will not break if the tension does not exceed the threshold given by C_{crit}. We now proceed to demonstrate that damage in the sense of the occurrence of large deformation gradients, takes place beyond this threshold. Before stating the precise results, we give an auxiliary statement which will be used in the proof of Theorem 3.8. We recall that a sequence of probability measures $\left\{\nu^k\right\}_{k\in\mathbb{N}}$ is *tight* if for each $\varepsilon > 0$, there exists a compact set K such that $\nu^k(K) > 1 - \varepsilon$ for every $k \in \mathbb{N}$.

Lemma 3.7 *Let $\{\nu^j\}_{j\in\mathbb{N}}$ be a sequence of solutions to the one-dimensional homogeneous time step problem* (17). *Assume that ν^j converges weakly-$\star$ to ν as $j \to \infty$ and that the sequence $\{\nu^j\}_{j\in\mathbb{N}}$ is tight. Then $E(\nu, \cdot)$ is minimised by ν.*

Proof: Let $\tilde{\nu}$ be a minimiser of $E(\nu, \cdot)$; suppose there is an $\eta > 0$ such that

$$E(\nu, \tilde{\nu}) \leq E(\nu, \nu) - \eta. \tag{23}$$

Since $\nu^j \stackrel{\star}{\rightharpoonup} \nu$ as $j \to \infty$ and the sequence $\{\nu^j\}_{j\in\mathbb{N}}$ is tight, it follows that $d_W^1\left(\nu^j, \nu\right) \to 0$ [38, Theorem 7.12]. By the triangle inequality, $d_W^1\left(\nu^j, \nu^{j+1}\right) \to d_W^1\left(\nu, \nu\right) = 0$. Furthermore, Lemma 2.2 is applicable and yields

$$\limsup_{j\to\infty} H_D^+\left(\operatorname{supp}\left(\nu^j\right); \operatorname{supp}\left(\tilde{\nu}\right)\right) \leq H_D^+\left(\operatorname{supp}\left(\nu\right); \operatorname{supp}\left(\tilde{\nu}\right)\right).$$

Thus,

$$\begin{aligned}
\limsup_{j\to\infty} E(\nu^j, \tilde{\nu}) &= \limsup_{j\to\infty} \left[\frac{h}{\varepsilon}\langle\phi, \tilde{\nu}\rangle + d_W^1\left(\nu^j, \tilde{\nu}\right)^2 + \delta^2 H_D^+\left(\operatorname{supp}\left(\nu^j\right); \operatorname{supp}\left(\tilde{\nu}\right)\right)^2\right] \\
&\leq \limsup_{j\to\infty} \left[\frac{h}{\varepsilon}\langle\phi, \tilde{\nu}\rangle + \left(d_W^1\left(\nu^j, \nu\right) + d_W^1\left(\nu, \tilde{\nu}\right)\right)^2\right] + \delta^2 H_D^+\left(\operatorname{supp}\left(\nu\right); \operatorname{supp}\left(\tilde{\nu}\right)\right)^2 \\
&= E(\nu, \tilde{\nu}).
\end{aligned} \tag{24}$$

Since ν^{j+1} is a minimiser for $E\left(\nu^j, \cdot\right)$ we have

$$E\left(\nu^j, \nu^{j+1}\right) \leq E\left(\nu^j, \nu^j\right) = \frac{h}{\varepsilon}\langle\phi, \nu^j\rangle. \tag{25}$$

Spelling out (25), we find that

$$\frac{h}{\varepsilon}\langle\phi, \nu^j\rangle \geq \frac{h}{\varepsilon}\langle\phi, \nu^{j+1}\rangle + d_W^1\left(\nu^j, \nu^{j+1}\right)^2 + \delta^2 H_D^+\left(\operatorname{supp}\left(\nu^j\right); \operatorname{supp}\left(\nu^{j+1}\right)\right)^2.$$

Thus,

$$\lim_{j\to\infty} \left[d_W^1\left(\nu^j, \nu^{j+1}\right)^2 + \delta^2 H_D^+\left(\operatorname{supp}\left(\nu^j\right); \operatorname{supp}\left(\nu^{j+1}\right)\right)^2\right] = 0$$

and

$$\begin{aligned}
E\left(\nu^j, \nu^{j+1}\right) &= \frac{h}{\varepsilon}\langle\phi, \nu^{j+1}\rangle + d_W^1\left(\nu^j, \nu^{j+1}\right)^2 + \delta^2 H_D^+\left(\operatorname{supp}\left(\nu^j\right); \operatorname{supp}\left(\nu^{j+1}\right)\right)^2 \\
&\to \frac{h}{\varepsilon}\langle\phi, \nu\rangle = E(\nu, \nu),
\end{aligned} \tag{26}$$

as $j \to \infty$. The combination of the two limits (24) and (26) with (23) shows that $E(\nu^j, \tilde{\nu}) < E\left(\nu^j, \nu^{j+1}\right)$ for j large enough, which contradicts the minimality of ν^{j+1} for $E(\nu^j, \cdot)$. □

Having investigated stability in Theorem 3.6, we are now in a position to focus on instability. The next theorem shows that for sufficiently large deformations, damage will occur. More precisely, it states that initial data with a deformation gradient $\nu^0 = \delta_{\tilde{A}}$ beyond the critical value C_{crit} (that is, $\tilde{A} > C_{\mathrm{crit}}$) are unstable with respect to small perturbations and lead inevitably to concentrations of the deformation gradient in Infinity as time goes to Infinity.

The next Theorem employs the concept of Young measure varifolds; we refer the reader to Appendix A for a short summary of this notion.

Theorem 3.8 (Instability) *If $A_j := \tilde{A} > C_{\mathrm{crit}}$, for every $j \in \mathbb{N}$, then the solution of* (11) *is unstable in the sense that if $\nu^0 \neq \delta_{\tilde{A}}$ and supp $\left(\nu^0\right) \subset (C_{\mathrm{crit}}, +\infty)$, then every subsequence of the solutions ν^j at time steps $j = 1, 2, \ldots$ converges for $j \to \infty$ to a Young measure varifold with a nontrivial varifold part.*

Proof: Since $\phi \geq 0$ is bounded from below by Assumption 2.4 (iii) we can deduce as in the proof of Theorem 3.6 that $\langle \phi, \nu^j \rangle$ is a monotonically decreasing sequence,

$$\langle \phi, \nu^j \rangle - \langle \phi, \nu^{j+1} \rangle \to 0 \text{ as } j \to \infty.$$

From this, we infer

$$\lim_{j \to \infty} d_W^1 \left(\nu^j, \nu^{j+1}\right) = 0$$

again as in the proof of Theorem 3.6. Since $\left\|\nu^j\right\| \leq 1$, there exists a subsequence (not relabelled) and a Young measure varifold Λ such that ν^j converges to Λ [18, Theorem 3.1]. Moreover, ν^j converges weak-$\star$ to a measure ν with $\|\nu\| \leq 1$. We want to show that $\|\nu\| < 1$, so let us suppose the opposite. We can distinguish two cases: either $\nu = \delta_{\tilde{A}}$ or $\nu \neq \delta_{\tilde{A}}$.

In the first case the definition of the functional E implies that, for j large enough, supp $\left(\nu^j\right) \subset (C_{\mathrm{crit}}, +\infty)$; we then say that the support of ν^j is entirely in the concave region of ϕ. An argument as in the proof of Lemma 3.5 then shows that $\left(\mathrm{supp}\left(\nu^j\right)\right)^{\mathrm{conv}}$ is a strict subset of $\left(\mathrm{supp}\left(\nu^{j+1}\right)\right)^{\mathrm{conv}}$. Thus ν^j cannot converge to ν.

Let us now suppose that $\|\nu\| = 1$ and $\nu \neq \delta_{\tilde{A}}$. Lemma 3.7 implies that a $\nu \neq \delta_{\tilde{A}}$ is a minimiser of $E(\nu, \cdot)$. We show that this cannot be the case: for any admissible Young measure ν, there exists an admissible Young measure $\tilde{\nu}$ such that $E(\nu, \tilde{\nu}) < E(\nu, \nu)$. The proof follows ideas similar to the proof of Theorem 3.6, hence we present the essential ideas and only sketch the precise computations.

We denote $b^j := \min \mathrm{supp}\left(\nu^j\right)$, $c^j := \max \mathrm{supp}\left(\nu^j\right)$ and S^j the convex envelope of the support of ν^j, that is,

$$S^j := \left[b^j, c^j\right] := \left(\mathrm{supp}\left(\nu^j\right)\right)^{\mathrm{conv}}.$$

Obviously $b^j < \tilde{A} < c^j$ holds, since it is assumed that $\nu^j \neq \delta_{\tilde{A}}$. A priori, $c^j = \infty$ is possible. We distinguish several cases.

Case 1: $\nu \llcorner (C_{\mathrm{crit}}, +\infty)$ is *not* concentrated in one point.

Choose μ with $\|\mu\| > 0$ and $\mu \leq \nu \llcorner (C_{\mathrm{crit}}, +\infty)$. We can apply the construction of Theorem 3.6, with $-\phi$ instead of ϕ, to μ in order to show that there is a measure $\hat{\mu}$ with $E(\mu, \hat{\mu}) < E(\mu, \mu)$. Therefore $\tilde{\nu} := \nu - \mu + \hat{\mu}$ satisfies $E(\nu, \tilde{\nu}) < E(\nu, \nu)$.

Case 2: $\nu \llcorner (C_{\mathrm{crit}}, +\infty)$ is concentrated in only one point, but $\nu \llcorner (0, C_{\mathrm{crit}}]$ is *not* concentrated in one point.

Obviously, $\nu \llcorner (0, C_{\mathrm{crit}}]$ cannot be the zero measure. (Otherwise, we would have $\nu = \delta_{\tilde{A}}$.) Hence this case can be handled analogously to the previous situation: We can directly apply the method from Theorem 3.6 on a submeasure μ with $\|\mu\| > 0$ and $\mu \leq \nu \llcorner (0, C_{\mathrm{crit}}]$ to obtain a Young measure $\tilde{\nu}$ with $E(\nu, \tilde{\nu}) < E(\nu, \nu)$.

Case 3: $\nu \llcorner (C_{\mathrm{crit}}, +\infty)$ and $\nu \llcorner (0, C_{\mathrm{crit}}]$ are each concentrated in one point, in other words, $\nu = (1-t)\delta_F + t\delta_G$ with $F \leq C_{\mathrm{crit}} < G$ and $t \in (0, 1)$.

We have to distinguish two sub-cases:

Case 3a: $\phi(F) \geq \phi(G) - \phi'(G)(G-F)$. This obviously implies

$$\phi'(G) \geq \frac{\phi(G)-\phi(F)}{G-F} \tag{27}$$

We claim that in this case

$$\phi(G) > \phi(F) + \phi'(F)(G-F). \tag{28}$$

Suppose the opposite inequality holds,

$$\phi'(F) \geq \frac{\phi(G)-\phi(F)}{G-F} \tag{29}$$

We remark that

$$\frac{\phi(G)-\phi(F)}{G-F} = \frac{1}{G-F}\int_F^G \phi'(X)\,\mathrm{d}X,$$

and $\phi'(F)$ and $\phi'(G)$ are both larger or equal than $\frac{\phi(G)-\phi(F)}{G-F}$ by (29) and (27), and ϕ' is not constant on (F,G). Therefore, since $\phi' \in C^1$, there must be a $Y \in (F,G)$ such that $\phi'(Y) < \frac{\phi(G)-\phi(F)}{G-F}$. This, however, implies that $\phi'' < 0$ somewhere on (F,Y) and $\phi'' > 0$ somewhere on (Y,G). Since $\phi'' > 0$ on $(0, C_{\mathrm{crit}})$ by Assumption 2.4 (ii) and $\phi'' < 0$ on $(C_{\mathrm{crit}}, +\infty)$ by Assumption 2.4 (iv), this is a contradiction. Hence, we have proved (28).

Let $\eta > 0$ and define $\nu_\eta := a(1-t)\delta_{F+\eta} + bt\delta_G$, where a and b are chosen such that the norm and expectation value of $\tilde{\nu}$ and ν coincide. With this information, we can compute a and b. Namely,

$$\|\nu_\eta\| = a(1-t) + bt.$$

Since both $\tilde{\nu}$ and ν are probability measures, we require $\|\nu_\eta\| = 1$, which holds if

$$b = \frac{1}{t} - a\frac{1-t}{t}.$$

A short computation shows that the equality of the expectation values $\langle \mathrm{Id}, \nu_\eta\rangle = a(1-t)(F+\eta) + btG$ and $\langle \mathrm{Id}, \nu\rangle = (1-t)F + tG$ amounts to

$$a = \frac{G-F}{G-F-\eta}.$$

We can now compare the energies of ν and ν_η.

$$\begin{aligned}
E(\nu,\nu_\eta) - E(\nu,\nu) &= \frac{h}{\varepsilon}\langle\phi,\nu_\eta\rangle - \frac{h}{\varepsilon}\langle\phi,\nu\rangle + d_W^1\,(\nu,\nu_\eta)^2 + \delta^2\eta^2 \\
&= \frac{h}{\varepsilon}\frac{G-F}{G-F-\eta}(1-t)\phi(F+\eta) + \frac{h}{\varepsilon}\left(\frac{1}{t} - \frac{(G-F)(1-t)}{(G-F-\eta)t}\right)t\phi(G) \\
&\quad - \frac{h}{\varepsilon}(1-t)\phi(F) - \frac{h}{\varepsilon}t\phi(G) \\
&\quad + ((a-1)(1-t)\eta + (1-b)t(G-F-\eta))^2 + \delta^2\eta^2 \\
&= \frac{h}{\varepsilon}(1-t)\frac{\eta}{G-F-\eta}\left(\phi'(F)(G-F) + \phi(F) - \phi(G)\right) + O(\eta^2).
\end{aligned}$$

By (28), the expression $\phi'(F)(G-F) + \phi(F) - \phi(G)$ is negative. Therefore, there exists an $\eta_0 > 0$ such that, for all $\eta \in (0,\eta_0]$, the energy difference is negative. With $\tilde{\nu} := \nu_{\eta_0}$, it thus follows that $E(\nu,\tilde{\nu}) < E(\nu,\nu)$, as claimed.

Case 3b: $\phi(F) < \phi(G) - \phi'(G)(G-F)$

In this case we define $\nu_\eta := a(1-t)\delta_F + bt\delta_{G+\eta}$ with a and b appropriately chosen. A computation similar to the previous case reveals that for small $\eta > 0$ we can choose $\tilde{\nu} := \nu_\eta$, and the conclusion $E(\nu,\tilde{\nu}) < E(\nu,\nu)$ holds as before. Hence, the theorem is proved. $\square$

One expects that in the quasi-static limit of the homogeneous case, damage will occur as soon as the expectation value A exceeds the critical value C_{crit}. This process is irreversible, since concentration corresponds to zero potential energy. Below this threshold, however, the behaviour can be described as elastic. This indicates that the model gives a reasonable description of the physical behaviour in a simplified context. It is expected that Truskinovsky's paradox [37] will not be observed in models involving a suitable dynamics, such as the model discussed here.

4 Existence and stability results for an alternative approximation

The time-continuous gradient flow (5) can also be approximated differently, inspired by minimising movements [2]. Analogues to some of the existence and stability results of the previous section are technically simpler in this setting and can even be proved in arbitrary space dimensions. It seems, however, much more difficult to study the convergence for a vanishing time step $h \to 0$. For the situation investigated in Section 3, but with a *symmetric* metric, this passage to the limit has been studied by Ambrosio et al. [4], whereas we are not aware of comparable results for the setting of this section. It thus seems that both approaches as presented in Section 3.2 and Section 4 are worth pursuing.

The idea of the approach in this section is to search at every time step for a state with minimal potential energy among all states close to the previous state (in a suitable metric). In the discretisation analysed in Section 3, closeness was indirectly enforced by the upper Hausdorff hemimetric.

The problem analysed in this section can be formulated as follows. Let

$$X := \{\nu \in \mathcal{G} \mid \langle \mathrm{Id}, \nu(x) \rangle = Du(x) \text{ for a.e. } x \in \Omega \text{ with } u \in W^{1,\infty}(\Omega, \mathbb{R}^n), u(x) = Ax \text{ on } \partial\Omega\}.$$

Let $\nu^0 \in X$ be a given gradient Young measure with compact support. Then ν^j is, for $j = 1, 2, \ldots$, determined as a minimiser of

$$E(\nu) := \int_\Omega \langle \phi, \nu \rangle \,\mathrm{d}x, \tag{30}$$

in X subject to the constraint

$$\int_\Omega d_W^1\left(\nu^j, \nu\right)\,\mathrm{d}x + \delta \sup_{x\in\Omega} d_H^+\left(\nu^j, \nu\right) \le \frac{h}{\varepsilon},$$

for given $h, \varepsilon > 0$. (We use a different scaling from the one in Section 3.)

In the following discussion, we again restrict ourselves to the spatially homogeneous situation.

The main result of this section is the following theorem. It proves existence of a solution for the model described by (30), and states the analogue of the stability and instability results of Section 3 (Theorem 3.6 and 3.8). It is shown that for initial data with small deformation gradients (below C_{crit}), elastic behaviour will occur, whereas for initial data with large deformation gradients (above C_{crit}), an instability leads to steadily growing deformation gradients, which corresponds to the formation of damage. We point out that for the model (30), the notion of stability is implicitly built into the model. Therefore, the following results are easier to formulate than the analogous results of Section 3. Again, the strong topology for Young measures is relevant; see (37) in Appendix A.

Theorem 4.1 (Stability II) *Consider the spatially homogeneous situation, where the measures ν^j and ν are constant as a function of $x \in \Omega$. Suppose $\operatorname{supp}\left(\nu^j\right)$ is bounded. Then, the time step problem* (30) *admits a solution, and for sufficiently small time steps $h > 0$ and fixed $\varepsilon > 0$, the following properties hold:*

1. If $\det(A) < C_{\text{crit}}$*, then a Dirac mass is stable: if* $\nu^j = \delta_A$*, then* $\nu^{j+1} = \delta_A$*.*

2. If $\det(A) > C_{\text{crit}}$*, then a Dirac mass is unstable: if* $\nu^j = \delta_A$*, then* $\nu^{j+1} \neq \delta_A$*.*

Proof: Existence can be shown as in the proof of Theorem 3.2; let

$$\mathcal{S} := \left\{ \nu \in \mathcal{G} \mid \langle \mathrm{Id}, \nu \rangle = A, \int_\Omega d_W^1 \left(\nu^j, \nu \right) \, \mathrm{d}x + \delta \sup_{x \in \Omega} d_H^+ \left(\nu^j, \nu \right) \leq \frac{h}{\varepsilon} \right\}.$$

First, we show that $\mathcal{S}$ is closed in the strong topology of Young measures (see Equation (37)). The set $\mathcal{S}$ is closed, since any sequence $\{\nu_k\}_{k \in \mathbb{N}} \subset \mathcal{S}$ of Young measures with $\nu_k \to \nu$ in the strong topology satisfies

$$\lim_{k \to \infty} \int_\Omega d_W^1 \left(\nu^j(x), \nu_k(x) \right) \, \mathrm{d}x = \int_\Omega d_W^1 \left(\nu^j(x), \nu(x) \right) \, \mathrm{d}x$$

and, by Lemma 2.2,

$$\lim_{k \to \infty} \sup_{x \in \Omega} d_H^+ \left(\nu^j(x), \nu_k(x) \right) \geq \sup_{x \in \Omega} d_H^+ \left(\nu^j(x), \nu(x) \right).$$

In addition, $\mathcal{S}$ is bounded, in the sense that the support of measures in $\mathcal{S}$ is uniformly bounded. Since for $p = 1$ (only), the space of Young measures with the p-Wasserstein metric is complete, we conclude that $\mathcal{S}$ is sequentially compact. Since E is bounded from below, a minimising sequence of measures ν_k with $\|\nu_k\| = 1$ exists. By compactness, there exists a converging subsequence of ν_k (not relabelled) and a limit measure $\nu \in \mathcal{S}$. (A standard diagonal argument shows that ν is a gradient Young measure.) Finally, sequential lower semicontinuity of E guarantees the existence of a minimiser for problem (30).

The first stability property follows from the fact that for $\frac{h}{\varepsilon} < C_{\text{crit}} - \det(A)$, within $B\left(A, \frac{h}{\varepsilon}\right)$, we can consider the quasiconvex function $\tilde{\phi}$ (Assumption 2.4 (ii)); compare Theorem A.3. To show the second property, choose a rank-one matrix C such that ϕ is strictly concave in the direction of C. This is possible by Assumption 2.4 (iv). One can again apply Jensen's inequality to conclude that $E\left(\lambda\left(\delta_{A-C}\right) + (1-\lambda)\left(\delta_{A+C}\right)\right) < E\left(\delta_A\right)$ for arbitrary $\lambda \in (0,1)$. By continuity, one can chose a $\lambda \in (0,1)$ such that the expectation value of $\lambda\left(\delta_{A-C}\right) + (1-\lambda)\left(\delta_{A+C}\right)$ is A. Moreover $\lambda\left(\delta_{A-C}\right) + (1-\lambda)\left(\delta_{A+C}\right)$ is by construction a gradient Young measure. Hence δ_A cannot be the minimiser of the time step problem (30). $\square$

The following inhomogeneous problem illustrates the need to work in the class of gradient Young measures, since the class of classical functions turns out to be too small.

Proposition 4.2 *Let* $\Omega \subset \mathbb{R}^2$ *be a bounded open domain with smooth boundary. Consider* $u \in W^{1,\infty}\left(\Omega, \mathbb{R}^2\right)$ *satisfying the boundary condition* $u(x) = Ax$ *with* $A \in \mathbb{R}^{2\times 2}$, $\det(A) > C_{\text{crit}}$*. Let* $\nu^0 = \delta_A$*. Then* ν^0 *is unstable, that is,* ν^1 *defined as solution of* (30) *satisfies* $\nu^1 \neq \nu^0$*. Moreover,* ν^1 *is a nontrivial gradient Young measure.*

Proof: By Theorem 4.1, $\nu^0 = \delta_A$ is unstable (in the sense defined there). Assume now that the minimiser $\tilde{\nu}$ of (30) is a trivial gradient Young measure, $\tilde{\nu}(x) = \delta_{Du(x)}$. Define ν as the homogenisation of $\tilde{\nu}$ by

$$\nu(E) := \frac{\left|\Omega \cap (Du)^{-1}(E)\right|}{|\Omega|}$$

for every Borel set $E \subset \mathbb{R}$.

We now prove that

$$\int_\Omega d_W^1 \left(\nu^0, \nu \right) \, \mathrm{d}x + \delta \sup_{x \in \Omega} d_H^+ \left(\nu^0, \nu \right) = \int_\Omega d_W^1 \left(\nu^0, \tilde{\nu} \right) \, \mathrm{d}x + \delta \sup_{x \in \Omega} d_H^+ \left(\nu^0, \tilde{\nu} \right). \tag{31}$$

To see this, first consider the Hausdorff metric part:

$$\begin{aligned}
\sup_x d_H^+\left(\nu^0, \nu\right) &= \sup_x d_H^+\left(\delta_A, \frac{\left|\Omega \cap (Du)^{-1}(\cdot)\right|}{|\Omega|}\right) \\
&= \sup_x H_D^+\left(\{A\}; \overline{Du(\Omega)}\right) \\
&= \sup_x \|A - Du(x)\| \\
&= \sup_x d_H^+\left(\nu^0, \tilde{\nu}\right).
\end{aligned}$$

To prove equality for the Wasserstein metric part in (31), we approximate $f := Du$ by step functions f_k, defined as

$$f_k(x) = \sum_{j=1}^{N_k} \chi_{Q_k^j}(x) f_k^j,$$

where $\left(Q_k^j\right)_j$ is a suitable finite decomposition of Ω with $\max_j \operatorname{diam}\left(Q_k^j\right) \to 0$ as $k \to \infty$, and

$$f_k^j := \frac{1}{\left|Q_k^j\right|} \int_{Q_k^j} f(x)\,\mathrm{d}x.$$

The corresponding homogenised measures are

$$\nu_k^j := \frac{\left|\Omega \cap (f_k\,|_{Q_k^j})^{-1}(\cdot)\right|}{|\Omega|} = \left|Q_k^j\right| \delta_{f_k^j}.$$

The sequence of measures $\nu_k := \sum_j \nu_k^j$ satisfies $\nu_k \overset{\star}{\rightharpoonup} \nu$. Therefore, we obtain

$$\begin{aligned}
d_W^1\left(\nu^0, \nu_k\right) &= \sum_j d_W^1\left(\left|Q_k^j\right| \delta_A, \left|Q_k^j\right| \delta_{f_k^j}\right) \\
&= \sum_j \left|Q_k^j\right| \cdot \left|A - f_k^j\right| \\
&\to \int_\Omega |A - Du(x)|\,\mathrm{d}x \text{ as } k \to \infty.
\end{aligned}$$

On the other hand, $d_W^1\left(\nu^0, \nu_k\right) \to d_W^1\left(\nu^0, \nu\right)$ as $k \to \infty$, and thus

$$\int_\Omega d_W^1\left(\nu^0, \nu\right)\,\mathrm{d}x = \int_\Omega |A - Du(x)|\,\mathrm{d}x = \int_\Omega d_W^1\left(\nu^0, \tilde{\nu}\right)\,\mathrm{d}x.$$

Taking everything together, one obtains a proof of (31).

It is now easy to see that $E(\nu^0, \nu) = E(\nu^0, \tilde{\nu})$. Thus, one can always find a nontrivial homogeneous gradient Young measure ν with the same energy as the given inhomogeneous measure $\tilde{\nu}$. In the class of homogeneous gradient Young measures, however, the minimum is not attained by ν, as can be seen by Jensen's inequality. Thus, the solution of the time step problem cannot be trivial, and ν_1 has to be a nontrivial gradient Young measure. □

In summary, some qualitative properties such as stability of the solution are easy to obtain for the approximation discussed in this section. However, the limit as $j \to \infty$ (that is, the asymptotic behaviour in time, which is particularly interesting for the quasistatic limit) of solutions in the unstable region is much more challenging.

5 Connections to other models

Let us relate the models under consideration to different recent approaches to fracture mechanics; we focus on the work by Francfort, Marigo [19], Dal Maso and Toader [12] and the concept of structured deformations invented by Del Piero and Owen [16, 14]. As a starting point, we recall Griffith's crack model [21]. Let $\Gamma \subset \Omega \subset \mathbb{R}^n$ denote a crack. A model for fracture is then given by the minimisation of

$$E(u, \Gamma) := \int_{\Omega \setminus \Gamma} \phi\left(Du(x)\right)\, \mathrm{d}x + \varepsilon \mathscr{H}^{n-1}(\Gamma),$$

where $\mathscr{H}^{n-1}$ denotes the $(n-1)$-dimensional Hausdorff measure, $\varepsilon > 0$ is constant, and suitable boundary conditions are applied. Here, the term $\mathscr{H}^{n-1}(\Gamma)$ acts as a penalisation for cracks and prevents immediate fracture. The model reflects Griffith's idea by assigning to a crack an energy density proportional to the surface of the crack. Griffith's model was refined by Francfort and Marigo [19], who consider a time-discretised evolution to describe crack formation and propagation. For a given initial condition with an initial crack Γ^0, one considers at every time step $j = 1, 2, \ldots$ the minimisation problem for $E(u, \Gamma)$ subject to the constraint $\Gamma^j \subset \Gamma$ (and some boundary conditions which might change in time).

The condition $\Gamma^j \subset \Gamma$ reflects the irreversibility of crack formation. Dal Maso and Toader [12] further modified the model by introducing a term $\left\|u^j - u\right\|^2$ to prevent rapid changes. Thus, they study the functional

$$E^j(u, \Gamma) := \int_{\Omega \setminus \Gamma} \left[\phi\left(Du(x)\right) + \left|u^j(x) - u(x)\right|^2\right] \mathrm{d}x + \varepsilon \mathscr{H}^{N-1}(\Gamma). \tag{32}$$

This functional is also known as the Mumford-Shah functional, introduced for edge detection in image decomposition [31]. Dal Maso and Toader [12] consider the limit $\varepsilon \to 0$ of (32) which yields a quasi-static model. Numerical simulations exhibit interesting examples of crack formation and propagation; compare also the work by Bourdin, Francfort and Marigo [6].

A key difference between the model (32) and the models studied in Sections 3 and 4 is that the latter ones do not include a surface energy. To understand the relation of the model proposed by Dal Maso and Toader and the ones discussed in this paper, it is thus natural to study the behaviour of (32) as $\varepsilon \to 0$. This problem has been studied in the context of edge detection and hence only for the two-dimensional scalar-valued case with $\phi(F) := |F|^2$ by Rieger and Tilli [35]. Before stating the result in this framework, we recall the definition of Γ-convergence; a detailed presentation can be found in the books by Braides [7] and Dal Maso [13].

Definition 5.1 (Γ-convergence) *Let $\{E_n\}_{n\in\mathbb{N}}$ be a sequence of functionals defined on a real Banach space X. Then E_n Γ-converges in X to the functional E, denoted $\Gamma\text{–}\lim_{n\to\infty} E_n = E$ (or $E_n \xrightarrow{\Gamma} E$ as $n \to \infty$), if the following two conditions hold true:*

1. *For every $u \in X$ and for every $u_n \to u$ in X we have*

$$\liminf_{n\to\infty} E_n(u_n) \geq E(u).$$

2. *For every $u \in X$, there exists a sequence $\{u_n\}_{n\in\mathbb{N}} \subset X$ such that $u_n \to u$ and*

$$\limsup_{n\to\infty} E_n(u_n) \leq E(u).$$

Motivated by the functional (32), we now wish to find a compact set $\Gamma \subset \Omega \subset \mathbb{R}^2$ which, for given $g \in H^1(\Omega)$, minimises the functional

$$\Gamma \mapsto \varepsilon \mathscr{H}^1(\Gamma) + \inf_{u \in H^1(\Omega\setminus\Gamma)} \int_{\Omega\setminus\Gamma} \left[|Du(x)|^2 + |u(x) - g(x)|^2\right] \mathrm{d}x. \tag{33}$$

We consider, for $\varepsilon > 0$, the family of rescaled functionals

$$J_\varepsilon(\Gamma) := \varepsilon \mathscr{H}^1(\Gamma) + \frac{1}{\varepsilon^2} \inf_{u \in H^1(\Omega\setminus\Gamma)} \int_{\Omega\setminus\Gamma} \left[|Du(x)|^2 + |u(x) - g(x)|^2 \right] \mathrm{d}x. \tag{34}$$

We choose a sequence $\varepsilon_n \to 0$ as $n \to \infty$. Since for small ε the integral term dominates and the contribution of the Hausdorff measure of the set Γ is of lower order, sequences of minimisers Γ_ε fill in general the entire domain Ω as $\varepsilon_n \to 0$. Thus it is natural to extend the domain of the functionals J_ε from sets to probability measures by identifying a non-empty compact set Γ of finite Hausdorff measure with the uniform probability measure given by $\mathscr{H}^1(\Gamma)^{-1}\mathscr{H}^1 \llcorner \Gamma$. More precisely, for probability measures μ on $\overline{\Omega}$, we define

$$E_\varepsilon(\mu) := \begin{cases} J_\varepsilon(\Gamma) & \text{if } \mu = \mathscr{H}^1(\Gamma)^{-1}\mathscr{H}^1 \llcorner \Gamma \text{ for some compact } \Gamma \subset \overline{\Omega} \\ & \text{with } 0 < \mathscr{H}^1(\Gamma) < \infty, \\ +\infty & \text{otherwise.} \end{cases} \tag{35}$$

We now state the main result of Tilli and Rieger [35], where the probability measure can be interpreted as the asymptotic probability of a crack being present at a given point.

Theorem 5.2 *Let $\Omega \subset \mathbb{R}^2$ be a bounded open set with Lipschitz boundary, and let $g \in H^1(\Omega)$. Consider the functional E_0 on probability measures in $\overline{\Omega}$ defined by*

$$E_0(\mu) := \left(\frac{9}{16}\right)^{\frac{1}{3}} \left(\int_\Omega \frac{|Dg(x)|^2}{\rho(x)^2} \,\mathrm{d}x \right)^{\frac{1}{3}}, \tag{36}$$

where $\rho \in L^1(\Omega)$ is the density of the absolutely continuous part of μ with respect to the Lebesgue measure.

Then the functionals E_ε defined in (35) *Γ-converge for $\varepsilon \to 0$ to the functional E_0, with respect to the weak-$\star$ topology on probability measures in $\overline{\Omega}$.*

This result can be applied accordingly to the functional (32) for crack propagation. An explicit computation of the minimiser of (36) gives the following result.

Proposition 5.3 *For the limit $\varepsilon \to 0$, the asymptotic probability of a crack at a point $x \in \Omega \subset \mathbb{R}^2$ in the model* (32) *proposed by Dal Maso and Toader, with $\phi(F) := |F|^2$, where the limit is taken as in Theorem 5.2, is*

$$p_j(x) := \frac{\left|Du^j(x)\right|^{\frac{2}{3}}}{\int_{\Omega\setminus\Gamma^j} |Du^j(y)|^{\frac{2}{3}} \,\mathrm{d}y}.$$

Thus, a crack is most likely to occur for maximal values of $\left|Du^j(x)\right|$.

If one compares this proposition with the results for the models proposed in this article, then one notices a similarity. Larger values of $\left|Du^j\right|$, respectively $|F|$, lead to a larger probability for fracture, respectively damage. By Proposition 5.3, the probability of fracture in the Γ-limit of a suitably scaled version of (32) increases with an increasing deformation. For the models discussed in the previous sections, damage occurs after a certain threshold is exceeded.

We also wish to sketch the connection of the models under consideration to the work of Del Piero, Owen and coworkers on structured deformations [14]. Structured deformations are a clever concept to describe fracture phenomena. The starting point are *simple deformations*, which are pairs (Γ, f), where $\Gamma \subset \Omega$ is a set of measure 0, and f is the *transplacement* defined on $\Omega \setminus \Gamma$ (see [15, Definition 3.2] for the precise definition). Thus, though regularity assumptions are imposed on f, the macroscopic transplacement may be discontinuous along Γ. The composition of simple deformations is then defined in such a way that the locus of the discontinuities can only increase under composition. *Structured deformations* are simple deformations augmented by a tensor field defined on $\Omega \setminus \Gamma$, subject to appropriate regularity assumptions (see [15, Definition

5.1]). In applications, the tensor field will be the gradient of the macroscopic transplacement; this needs to be introduced since for a sequence of simple transformations, the limit of the gradient does not necessarily coincide with the gradient of the limit. It can be shown that every structured deformation is the limit of simple deformations [15, Theorem 5.8]. Both the models under consideration and structured deformations have in common that the irreversibility is built into the model; however, while this is here a consequence of the asymmetry of the metric, it follows for structured deformations — perhaps more elegantly — via the composition of maps inherited from simple deformations. In this article, the information pertaining to the tensor of structured deformations is encoded in the Young measure varifold via the expectation value. However, the concept of structured deformations does not rely on the support of the Young measure, which plays a crucial rôle in the model analysed here. Structured deformations can be used to describe fracture phenomena. For example, for a model where the elastic bulk energy is augmented by a suitable interfacial energy, it can be shown that a transition from ductile to brittle fracture takes place as the length of the reference bar increases [10]. For further information on structured deformations, we refer to [15, 16, 14].

6 Discussion

The starting point for this work is the observation that many problems in nature, including fracture and damage, do not seem to attain the ground state of the energy. We discussed two time-discretised models of gradient flows for Young measures and showed that these models exhibit in the homogeneous, one-dimensional situation a more realistic picture than global energy minimisation.

The models investigated here are phenomenological, but introduce no parameters other than those necessary to describe a non-quasi-convex energy density. They are a natural extension of models of classical elasticity in the sense that they agree with those of classical elasticity for deformation gradients below the threshold at which the energy becomes concave. Related models have been proposed for rate-independent models of plasticity and phase transformations [28, 26]. It is remarkable that in both situations, asymmetric metrics appear rather naturally. In Section 5, connections to further models have been given, notably to the model of Dal Maso and Toader [12] for fracture, and to structured deformations as developed by Del Piero and Owen [16, 14].

A Young measures and varifolds

Young measures have been successfully applied in various problems in mathematical materials science. For example, they describe microstructures arising in martensitic materials. This approach has been pioneered by Ball and James [5, 29], where Young measures describe oscillations in microstructured materials modelled by a nonconvex free energy density. Young measure varifolds provide an extension to concentration phenomena [18] which arise, for example, in situations where the energy density has a sublinear growth at infinity. For the reader's convenience, some information on Young measures and Young measure varifolds is collated in this Appendix.

A *Young measure* [39] (or *parameterised measure*) is a weakly-$\star$ measurable mapping

$$\Omega \to \mathrm{Prob}\left(\mathbb{R}^m\right), \qquad x \mapsto \nu_x$$

with values in the probability measures. (We recall that the mapping ν is *weakly-$\star$ measurable* if for any $\phi \in C_0\left(\mathbb{R}^m, \mathbb{R}\right)$, the mapping

$$\Omega \to \mathbb{R}, \qquad x \mapsto \overline{\phi}(x) := \langle \phi, \nu_x \rangle := \int_{\mathbb{R}^m} \phi(F)\, \mathrm{d}\nu_x(F)$$

is measurable in the usual sense.) A *Young measure with finite pth moment* is a weakly-$\star$ measurable map as above, but with values in the set $\mathrm{Prob}^p\left(\mathbb{R}^m\right)$ of probability measures with finite pth moment, see the definition after Equation (2).

A sequence of Young measures $\{\nu_x^j\}_{j\in\mathbb{N}}$ is said to converge weakly-$\star$ to a Young measure ν_x if for every $\theta \in C_0(\Omega, \mathbb{R})$ and every $\phi \in C_0(\mathbb{R}^m, \mathbb{R})$ one has

$$\int_\Omega \theta(x) \int_{\mathbb{R}^m} \phi(F)\, \mathrm{d}\nu_x^j(F)\, \mathrm{d}x \to \int_\Omega \theta(x) \int_{\mathbb{R}^m} \phi(F)\, \mathrm{d}\nu_x(F)\, \mathrm{d}x \text{ as } j \to \infty.$$

There is another metric for Young measures defined by the Wasserstein distance for probability measures, see Equation (9). Namely, the *Wasserstein metric for Young measures* is given by

$$\mathrm{dist}^1(\nu, \mu) := \int_\Omega d_W^1(\nu(x), \mu(x))\, \mathrm{d}x. \tag{37}$$

Convergence in this metric implies weak-$\star$ convergence, but not vice versa [25]. Also, when regarding classical functions f, g as Young measures, one can see that

$$\mathrm{dist}^1(\delta_f, \delta_g) = \|f - g\|_{L^1(\Omega)}. \tag{38}$$

A $W^{1,\infty}$*-gradient Young measure* is a Young measure associated with sequences $\{f_j\}_{j\in\mathbb{N}}$ of gradients of $W^{1,\infty}$-functions. In the one-dimensional case, the classes of Young measures and gradient Young measures coincide. For a given domain $\Omega \subset \mathbb{R}^n$ and a fixed target space $\mathbb{R}^m$, the set of $W^{1,\infty}$-gradient Young measures with finite first moment is denoted $\mathcal{G}$. For simplicity we usually refer to them as *Young measures*. To highlight the interpretation as a map and to avoid misunderstandings with the notation for derivatives, we often write $\nu(x)$ for ν_x.

Let us recall the notion of Young measure varifolds [18]. We denote the unit sphere in $\mathbb{R}^d$ by S^{d-1}.

Definition A.1 (Young measure varifolds) *Let $1 < q < \infty$. A pair consisting of a gradient Young measure $\nu\colon \Omega \to \mathcal{M}$ and a nonnegative Radon measure Λ on $\Omega \times S^{mn-1}$ is a* Young measure varifold *if there exists a sequence $\{u_j\}_{j\in\mathbb{N}}$ of functions $u_j \in W^{1,\infty}(\Omega, \mathbb{R}^m)$ such that for every $\theta \in C_0(\Omega, \mathbb{R})$, every $\phi \in C(\mathbb{R}^{m\times n}, \mathbb{R})$ with $|\phi(\xi)| \le C(1 + |\xi|^r)$ for $1 \le r < q$ and for every q-homogeneous $\psi \in C(\mathbb{R}^{m\times n}, \mathbb{R})$, the following statements hold.*

$$\int_\Omega \theta(x)\phi(Du_j(x))\, \mathrm{d}x \to \int_\Omega \theta(x) \int_{\mathbb{R}^{m\times n}} \phi(F)\, \mathrm{d}\nu(x)(F)\, \mathrm{d}x, \tag{39}$$

$$\int_\Omega \theta(x)\psi(Du_j(x))\, \mathrm{d}x \to \int_{\Omega\times S^{mn-1}} \theta(x)\psi(F)\, \mathrm{d}\Lambda(x, F). \tag{40}$$

Using the slicing decomposition [17, Section 1.E] for Λ, we can write $\Lambda = \lambda \otimes \pi$, where π is the projection of Λ onto Ω. Hence, (40) can be restated as

$$\int_\Omega \theta(x)\psi(Du_j(x))\, \mathrm{d}x \to \int_\Omega \theta(x) \int_{S^{mn-1}} \psi(F)\, \mathrm{d}\lambda(x)(F)\, \mathrm{d}\pi(x).$$

The intuition behind this definition is that the limit measure of a sequence of functions might not be of unit mass if the sequence transports mass "to Infinity". The varifold Λ measures how much mass is leaking out and records the direction. An alternative way of understanding this phenomenon is to consider the compactification of $\mathbb{R}^{mn}$ by S^{mn-1}. Young measures defined on the compact space preserve mass in the limit. It can be seen that the two different viewpoints result in the same object, namely Young measure varifolds. More information on Young measure varifolds can be found elsewhere [1, 18].

In this article, Young measure varifolds appear only as limiting objects as time goes to Infinity. Young measures describe classical deformation gradients (that is, functions of x) and possible microstructures. If one interprets damage as infinitely large deformation gradients, then a non-vanishing varifold part describes damage (in engineering applications, one will be inclined to say that damage occurs where deformation gradients exceed a certain threshold; yet, varifolds are an appropriate concept to deal with the possibility of ever larger deformation gradients in the limit as time goes to Infinity). Should the set of points in Ω with non-vanishing varifold part form a low-dimensional subset of Ω, then one can say that the material damage occurs in form of fracture.

We recall the following generalisation of convexity.

Definition A.2 *A function $f\colon \mathbb{R}^{m\times n}\to\mathbb{R}$ is* quasiconvex *if for every open and bounded set U with $|\partial U|=0$ one has for every $\zeta\in W_0^{1,\infty}(U,\mathbb{R}^m)$*

$$\int_U f(F+D\zeta)\,\mathrm{d}x-|U|\,f(F)\geq 0,$$

whenever the integral on the left-hand side exists.

A function f is called uniformly quasiconvex *if there exists a positive constant C such that*

$$\int_U f(F+D\zeta)\,\mathrm{d}x-|U|\,f(F)\geq C\int_U |D\zeta|^2\,\mathrm{d}x,$$

Every convex function is quasiconvex, and if $n=1$ or $m=1$, both notions coincide.

The following characterisation of gradient Young measures has been proved by Kinderlehrer and Pedregal [23].

Theorem A.3 *Let $\nu\colon\Omega\to\mathcal{M}(\mathbb{R}^{m\times n})$ be a weak-$\star$ measurable map. Then ν is a gradient Young measure if and only if $\nu(x)\geq 0$ a.e. and there exists a compact set $K\subset\mathbb{R}^{m\times n}$ and $u\in W^{1,\infty}(\Omega,\mathbb{R}^m)$ such that the following three conditions hold.*

(i) supp $(\nu(x))\subset K$ for a.e. $x\in\Omega$,

(ii) $\langle\mathrm{Id},\nu(x)\rangle=Du(x)$ for a.e. $x\in\Omega$,

(iii) $\langle f,\nu(x)\rangle\geq f(\langle\mathrm{Id},\nu(x)\rangle)$ for a.e. $x\in\Omega$ and every quasiconvex function $f\colon\mathbb{R}^{m\times n}\to\mathbb{R}$.

Acknowledgements

This work started while MOR was a postdoctoral scholar at Carnegie Mellon University and JZ at the California Institute of Technology. Both authors are grateful to Kaushik Bhattacharya and Alexander Mielke for extensive discussions. MOR thanks the Scuola Normale Superiore and the NCCR FINRISK, Project 3, for their financial support, and the University of Bath and the Max-Planck-Institute for Mathematics in the Sciences, Leipzig, for their hospitality during several visits. JZ gratefully acknowledges the financial support of the US Air Force Office of Scientific Research through an NSF-ITR grant (ACI-0204932), the Deutsche Forschungsgemeinschaft through an Emmy Noether grant (Zi 751/1-1) and the EPSRC through an Advanced Research Fellowship (GR / S99037 / 1). Both authors acknowledge financial support of the EPSRC (GR / S99044 / 1). We thank Christoph Ortner and a referee for valuable suggestions and Julia Collins for helpful comments; Kaushik Dayal kindly pointed out some relevant literature.

Marc Oliver Rieger, ISB, Universität Zürich, Switzerland, rieger@isb.uzh.ch

Johannes Zimmer, University of Bath, Department of Mathematical Sciences, Bath BA2 7AY, United Kingdom

References

[1] J. J. Alibert and G. Bouchitté. Non-uniform integrability and generalized Young measures. *J. Convex Anal.*, 4(1):129–147, 1997.

[2] Luigi Ambrosio. Minimizing movements. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5)*, 19:191–246, 1995.

[3] Luigi Ambrosio. Lecture notes on optimal transport problems. In P. Colli and C.F. Rodrigues, editors, *CIME summer school in Madeira, Lecture Notes in Mathematics*, volume 1813, pages 1–52. Springer Verlag, 2003.

[4] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[5] J. M. Ball and R. D. James. Fine phase mixtures as minimizers of energy. *Arch. Rational Mech. Anal.*, 100(1):13–52, 1987.

[6] B. Bourdin, G. A. Francfort, and J.-J. Marigo. Numerical experiments in revisited brittle fracture. *J. Mech. Phys. Solids*, 48(4):797–826, 2000.

[7] Andrea Braides. *Γ-convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[8] Y. P. Chen, J. D. Lee, and A. Eskandarian. Atomistic viewpoint of the applicability of microcontinuum theories. *International Journal of Solids and Structures*, 41(8):2085–2097, 2004. 19.

[9] Isaac Chenchiah, Marc Oliver Rieger, and Johannes Zimmer. Gradient flows in asymmetric metric spaces. In preparation.

[10] R. Choksi, G. Del Piero, I. Fonseca, and D. Owen. Structured deformations as energy minimizers in models of fracture and hysteresis. *Math. Mech. Solids*, 4(3):321–356, 1999.

[11] G. Dal Maso, A. DeSimone, M. G. Mora, and M. Morini. Time-dependent systems of generalized Young measures. *Netw. Heterog. Media*, 2(1):1–36 (electronic), 2007.

[12] G. Dal Maso and R. Toader. A model for the quasi-static growth of brittle fractures based on local minimization. *Math. Models Methods Appl. Sci.*, 12(12):1773–1799, 2002.

[13] Gianni Dal Maso. *An introduction to Γ-convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[14] Gianpietro Del Piero. One-dimensional ductile-brittle transition, yielding, and structured deformations. In *Variations of domain and free-boundary problems in solid mechanics (Paris, 1997)*, volume 66 of *Solid Mech. Appl.*, pages 203–210. Kluwer Acad. Publ., Dordrecht, 1999.

[15] Gianpietro Del Piero and David R. Owen. Structured deformations of continua. *Arch. Rational Mech. Anal.*, 124(2):99–155, 1993.

[16] Gianpietro Del Piero and David R. Owen. Integral-gradient formulae for structured deformations. *Arch. Rational Mech. Anal.*, 131(2):121–138, 1995.

[17] Lawrence C. Evans. *Weak convergence methods for nonlinear partial differential equations.* American Mathematical Society, Providence, R.I., 1990.

[18] Irene Fonseca, Stefan Müller, and Pablo Pedregal. Analysis of concentration and oscillation effects generated by gradients. *SIAM J. Math. Anal.*, 29(3):736–756 (electronic), 1998.

[19] G. A. Francfort and J.-J. Marigo. Revisiting brittle fracture as an energy minimization problem. *J. Mech. Phys. Solids*, 46(8):1319–1342, 1998.

[20] G. Friesecke and J. B. McLeod. Dynamics as a mechanism preventing the formation of finer and finer microstructure. *Arch. Rational Mech. Anal.*, 133(3):199–247, 1996.

[21] A. Griffith. The penomena of rupture and flow in solids. *Philos. Trans. R. Soc. London, Ser. A*, 221:163–198, 1920.

[22] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17 (electronic), 1998.

[23] David Kinderlehrer and Pablo Pedregal. Characterizations of Young measures generated by gradients. *Arch. Rational Mech. Anal.*, 115(4):329–365, 1991.

[24] Isaak A. Kunin. *Elastic media with microstructure. I*, volume 26 of *Springer Series in Solid-State Sciences*. Springer-Verlag, Berlin, 1982. One-dimensional models, Translated from the Russian.

[25] Alexander Mielke. Flow properties for Young-measure solutions of semilinear hyperbolic problems. *Proc. Roy. Soc. Edinburgh Sect. A*, 129(1):85–123, 1999.

[26] Alexander Mielke. Deriving new evolution equations for microstructures via relaxation of variational incremental problems. *Comput. Methods Appl. Mech. Engrg.*, 193(48-51):5095–5127, 2004.

[27] Alexander Mielke. Evolution of rate-independent systems. In *Evolutionary equations. Vol. II*, Handb. Differ. Equ., pages 461–559. Elsevier/North-Holland, Amsterdam, 2005.

[28] Alexander Mielke and Tomáš Roubíček. A rate-independent model for inelastic behavior of shape-memory alloys. *Multiscale Model. Simul.*, 1(4):571–597 (electronic), 2003.

[29] Stefan Müller. Variational models for microstructure and phase transition. In S. Hildebrandt and M. Struwe, editors, *Calculus of Variations and Geometric Evolution Problems*, number 1713 in Lecture Notes in Mathematics, page 211 ff. Springer-Verlag, Berlin etc., 1999.

[30] Stefan Müller. Variational models for microstructure and phase transitions. In *Calculus of variations and geometric evolution problems (Cetraro, 1996)*, volume 1713 of *Lecture Notes in Math.*, pages 85–210. Springer, Berlin, 1999.

[31] David Mumford and Jayant Shah. Boundary detection by minimizing functionals. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition.* San Francisco, 1985.

[32] Svetlozar T. Rachev. *Probability metrics and the stability of stochastic models.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1991.

[33] Marc Oliver Rieger. A model for hysteresis in mechanics using local minimizers of Young measures. In *Progress in Nonlinear Differential Equations and Their Applications*, volume 63, pages 363–374. Birkhäuser Verlag Basel, 2005.

[34] Marc Oliver Rieger. Monotonicity of transport plans and applications. CVGMT preprint, Pisa, 2006.

[35] Marc Oliver Rieger and Paolo Tilli. On the Γ-limit of the Mumford-Shah functional. *Calculus of Variations and Partial Differential Equations*, 23(4):373–390, 2005.

[36] S. A. Silling. Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids*, 48(1):175–209, 2000.

[37] L. Truskinovsky. Fracture as a phase transformation. In R. C. Batra and M. F. Beatty, editors, *Contemporary research in mechanics and mathematics of materials*, pages 322–332. CIMNE, 1996.

[38] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2003.

[39] L. C. Young. Generalized curves and the existence of an attained absolute minimium in the calculus of variations. *Comptes Rendus de la Soc. des Sciences et de Lettres de Varsovie, classe III*, 30:212–34, 1937.